ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

TOMUS 50

FASC. 1—2

SZEGED, 1986

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNYEI

ACTA SCIENTIARUM MATHEMATICARUM

50. KÖTET

FASC. 1—2

SZEGED, 1986

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI INTÉZETE

ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

TOMUS 50

SZEGED, 1986

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNYEI

ACTA SCIENTIARUM MATHEMATICARUM

50. KÖTET

SZEGED, 1986

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI INTÉZETE

$\boxed{\text{András P. Huhn}}$

(1947—1985)

It is a great loss to the Hungarian mathematical community, to the University of Szeged and to Acta Scientiarum Mathematicarum that András P. Huhn, a member of the editorial board of this Acta, suddenly passed away on June 6, 1985.

He was born in Szeged on January 26, 1947, and attended school here in his native town. Having completed his mathematical studies at the University of Szeged, in 1971, he started to work with the Research Group on Mathematical Logic and Automaton Theory of the Hungarian Academy of Sciences. In 1973, he came to the Department of Algebra and Number Theory of Szeged University, where he became an Associate Professor in 1977. He received his university doctorate with the distinction "Sub Auspiciis Rei Publicae Popularis" in 1973, acquired his Candidate's Degree in 1975, and had almost completed his thesis for his Degree of Doctor of the Academy when a tragic accident caused his decease.

András P. Huhn started his mathematical activities in 1969 with the theory of lattices, which became his main research field. Generalizing the notion of distributivity he introduced the concept of n-distributive lattices, i.e., lattices obeying the n-distributive law $x\wedge\bigvee_{i=0}^{n} y_i=\bigvee_{j=0}^{n}\left(x\wedge\bigvee_{\substack{i=0\\ i\neq j}}^{n} y_i\right)$. One of his main achievements in lattice theory is the discovery of the fact that a modular lattice is n-distributive if and only if it does not include a certain partial lattice, called n-diamond, as a sublattice. He found a great number of important applications for n-distributivity throughout lattice theory and universal algebra, the most remarkable being the realization that n-diamonds are equivalent to von Neumann's frames of order $n+1$, which brought n-distributivity into connection with von Neumann's coordinatization theory. More than half of his papers deal with n-distributivity and its applications. He obtained many other important results, e.g., on finitely presented lattices, free products and submodule lattices.

András P. Huhn always strove for thoroughness and, while being a sensitive, open-minded person in a general way, was fully devoted to his profession. Both colleagues and students liked him very much. He organized three successful mathematics conferences in Szeged and, besides his more than thirty mathematical papers, edited two colloquium proceedings on lattice theory. His mathematical activities brought him high international repute. He was invited to the editorial board of Algebra Universalis, the leading journal in his research field, and of our Acta Scientiarum Mathematicarum. His decease is a great loss for our community. We will cherish his memory.

Acta Sci. Math., 50 (1986), 5—14

Involution algebras and the Anderson—Divinsky—Suliński property

N. V. LOI and R. WIEGANDT

1. Introduction

In this paper we shall deal with special features of the general radical theory of involution algebras. We give necessary and sufficient conditions for a radical class $\mathbf{R}$ to have the A—D—S property:

$$I^* \lhd A^* \quad \text{implies} \quad \mathbf{R}(I^*) \lhd A^*$$

(where * indicates that algebras with involution * are considered). We prove that every radical class of involution algebras over a field K has the A—D—S property if and only if $\operatorname{char} K=2$. The A—D—S property implies trivially the hereditariness of the corresponding semisimple class. Nevertheless, there are radical classes which do not have the A—D—S property, but have hereditary semisimple classes. The semisimple class of a radical need not be hereditary, even if the radical is hereditary.

A K-algebra A is an involution algebra, if in A a unary operation * is defined so that $x^{**}=x$, $(x+y)^*=x^*+y^*$, $(xy)^*=y^*x^*$, $(kx)^*=kx^*$ for all $x, y \in A$ and $k \in K$. We shall work throughout with involution algebras over a commutative associative ring K with identity, and the universal class we shall use will be the variety $\mathfrak{V}$ of all K-algebras with involution. A^* will always stand for a K-algebra with involution *. In particular, id will denote the operation $x^{\mathrm{id}}=x$ and $^{-*}$ the operation $x^{-*}=-x^*$. An ideal I^* of an involution algebra A^* will always mean an ideal of the algebra A such that I^* is an involution algebra. This fact will be indicated by $I^* \lhd A^*$. By a homomorphism φ we mean an algebra homomorphism such that $(\varphi(x))^*=\varphi(x^*)$.

A *radical class* $\mathbf{R}$ (in the sense of Kurosh and Amitsur) of involution algebras is a subclass $\mathbf{R}$ of $\mathfrak{V}$ such that

i) $\mathbf{R}$ is homomorphically closed: if $A^* \in \mathbf{R}$ then $\varphi(A^*) \in \mathbf{R}$ for every homomorphism φ,

Received April 2, 1984.

ii) for every $A^*\in\mathfrak{V}$, the ideal $\mathbf{R}(A^*)=\sum(I^*\lhd A^*\colon I^*\in\mathbf{R})$ is in $\mathbf{R}$,

iii) $\mathbf{R}$ is closed under extensions: if $I^*\lhd A^*$, $I^*\in\mathbf{R}$ and $(A/I)^*\in\mathbf{R}$, then $A^*\in\mathbf{R}$.

Condition iii) can be replaced by $\mathbf{R}(A^*/\mathbf{R}(A^*))=0$ for all $A^*\in\mathfrak{V}$. The class

$$\mathcal{S}\mathbf{R}=\{A^*\in\mathfrak{V}\colon \mathbf{R}(A^*)=0\}$$

is called the *semisimple class* of the radical class $\mathbf{R}$. A semisimple class $\mathbf{S}$ is always *regular,* that is, if $0\neq I^*\lhd A^*\in\mathbf{S}$, then there exists an $M^*\lhd I^*$ such that $0\neq(I/M)^*\in\mathbf{S}$. If $\mathbf{M}$ is any regular class, in particular a semisimple class, then the class

$$\mathcal{U}\mathbf{M}=\{A^*\in\mathfrak{V}\colon A^*/I^*\in\mathbf{M}\Rightarrow I^*=A^*\}$$

is a radical class, which is referred to as the *upper radical class* of $\mathbf{M}$. For further details of the basic facts of radical theory we refer to [9]. Radicals of involution algebras have been studied in the recent papers [3], [5], [6] and [8].

Given a radical class $\mathbf{R}$, it may happen that

$$I^*\lhd A^*\in\mathfrak{V}\quad\text{implies}\quad \mathbf{R}(I^*)\lhd A^*.$$

In this case we say that $\mathbf{R}$ *satisfies A—D—S.* If $\mathbf{R}$ satisfies A—D—S, then it follows trivially that

$$I^*\lhd A^*\in\mathfrak{V}\quad\text{implies}\quad \mathbf{R}(I^*)\subseteq\mathbf{R}(A^*).$$

This latter condition is equivalent to demanding that

$$I^*\lhd A^*\in\mathcal{S}\mathbf{R}\quad\text{implies}\quad I^*\in\mathcal{S}\mathbf{R},$$

that is, that the semisimple class $\mathcal{S}\mathbf{R}$ is *hereditary.* Every radical class of algebras (without involution) satisfies A—D—S; this statement is the Anderson—Divinsky—Suliński Theorem [1], which is of fundamental importance in the general theory of radicals. For involution algebras, however, this is not always so, SALAVOVÁ ([5] Example 1, 9) gave a radical class $\mathbf{R}$ whose semisimple class is not hereditary, and consequently $\mathbf{R}$ does not satisfy A—D—S. In [8] WICHMANN suggested to deal with the problem whether a given radical class $\mathbf{R}$ of involution algebras satisfies A—D—S. This and related questions will be the topic of our investigations.

2. The A—D—S property

The next propositions show us that the variety $\mathfrak{V}$ of involution algebras is quite a good one.

Proposition 1. *If* $I^*\lhd A^*$, *then* $(I^2)^*\lhd A^*$.

Proof. It is clear that I^2 is an ideal of the algebra A. If $a, b \in I$, then

$$(ab)^* = b^* a^* \in I^2,$$

and so the assertion holds by the additivity of *.

Proposition 2. *For involution algebras the Andrunakievich Lemma holds; if* $M^* \lhd I^* \lhd A^*$ *and* L^* *denotes the ideal of* A^* *generated by* M^*, *then* $(L^*)^3 \subseteq M^*$.

Proof. We shall prove that the ideal J of the algebra A (without involution) generated by M is closed under involution. If $d \in J$, then

$$d = m + \sum_i a_i x_i + \sum_j y_j b_j + \sum_l u_l c_l v_l$$

where m, a_i, b_j, c_l are in M and x_i, y_j, u_l, v_l are in A. Since

$$d^* = m^* + \sum_i x_i^* a_i^* + \sum_j b_j^* y_j^* + \sum_l v_l^* c_l^* u_l^* \in M + AM + MA + AMA = J,$$

it follows that J^* is an involution algebra. Thus the Andrunakievich Lemma for algebras infers the assertion.

In view of Propositions 1 and 2, the proof of [2] Theorem 3, 2 and [2] Proposition 3, 5 yield immediately the following (see also [5] Предложение 3.4).

Corollary 1. *If* $\mathbf{R}$ *is a radical class of involution algebras which either contains all involution algebras with zero-multiplication or consists of idempotent involution algebras, then* $\mathbf{R}$ *satisfies A—D—S and hence the semisimple class* $\mathcal{S}\mathbf{R}$ *is hereditary.*

In order to prove necessary and sufficient conditions for a radical class $\mathbf{R}$ to satisfy A—D—S, we have to develop some techniques.

Proposition 3. *Let* A^* *be an involution algebra such that* $A^2=0$. *Then* A^{-*} *is also an involution algebra.*

Proof. Obvious.

Proposition 4. *Let* A^* *be an involution algebra such that* $A^2=0$. *If the unary operation* $^\square$ *of* $B=A\oplus A$ *is defined by* $(x,y)^\square=(y^*,x^*)$, *then* $B^\square$ *is an involution algebra.*

Proof. Straightforward.

Proposition 5. *Let* $\mathbf{R}$ *be a radical class of involution algebras,* $I^* \lhd A^*$ *and* $L^*=\mathbf{R}(I^*)$. *If* $(L^2/L^3)^* \in \mathbf{R}$, *then for any element* $a \in A$,

(i) $aL^2a^* \subseteq L$,

(ii) $(aL+La^*+L)^* \lhd I^*$,

(iii) *the mapping*

$$\varphi_a\colon ((L/L^2)\oplus(L/L^2))^\square \to ((aL+La^*+L)/L)^*$$

defined by $\varphi_a(x+L^2, y+L^2)=ax+ya^*+L$ *is a homomorphism onto*

$$((aL+La^*+L)/L)^*.$$

Proof. Let us define the mapping

$$\psi\colon (L^2/L^3)^* \to ((aL^2a^*+L)/L)^*$$

by $\psi(\sum x_iy_i+L^3)=a(\sum x_iy_i)a^*+L$. It is easy to check that ψ is a homomorphism onto $((aL^2a^*+L)/L)^*$, hence by $(L^2/L^3)^*\in\mathbf{R}$ we have

$$((aL^2a^*+L)/L)^*\in\mathbf{R}.$$

Since $((aL^2a^*+L)/L)^*\lhd I^*/L^*$, it follows

$$((aL^2a^*+L)/L)^* \subseteq \mathbf{R}(I^*/L^*) = \mathbf{R}(I^*/\mathbf{R}(I^*)) = 0.$$

Hence $aL^2a^*\subseteq L$ holds proving (i).

Using (i), one can verify easily that

$$(aL+La^*+L)^*\lhd I^*.$$

The last assertion is a straightforward calculation with an application of Proposition 4.

Proposition 6. *Let* $\mathbf{R}$ *be a radical class of involution algebras satisfying the following condition:*

(*) *if* $A^*\in\mathbf{R}$ *and* $A^2=0$, *then* $A^\circ\in\mathbf{R}$ *for any involution* $^\circ$ *on* A.
Then for any $L^*\in\mathbf{R}$, *also* $(L^2/L^3)^*\in\mathbf{R}$.

Proof. Let us consider an arbitrary element $a\in L$ and the mapping

$$f_a\colon L/L^2 \to L^2/L^3$$

defined by $f_a(x+L^2)=ax+L^3$ for all $x\in L$. The mapping f_a is obviously a homomorphism of the algebra $(L/L^2)^{\mathrm{id}}$ into $(L^2/L^3)^{\mathrm{id}}$ and $f_a(L/L^2)^{\mathrm{id}}\lhd(L^2/L^3)^{\mathrm{id}}$. Since $L^*\in\mathbf{R}$, condition (*) implies $(L/L^2)^{\mathrm{id}}\in\mathbf{R}$. Thus also $f_a(L/L^2)^{\mathrm{id}}\in\mathbf{R}$. Hence by

$$(L^2/L^3)^{\mathrm{id}} = \sum_{a\in L} f_a(L/L^2)^{\mathrm{id}} \subseteq \mathbf{R}((L^2/L^3)^{\mathrm{id}}) \subseteq (L^2/L^3)^{\mathrm{id}}$$

we have $(L^2/L^3)^{\mathrm{id}}\in\mathbf{R}$ and so condition (*) yields $(L^2/L^3)^*\in\mathbf{R}$.

Let A^* be an involution algebra over the ring K such that $A^2=0$. The ring K can be regarded as an involution algebra K^{id}. Let us consider the Cartesian product $E=K\times K\times A\times A$. On E we define operations by the following rules:

$$(a,b,x,y)+(c,d,u,v) = (a+c, b+d, x+u, y+v),$$

$$(a,b,x,y)(c,d,u,v) = (ac-bd, ad+bc, au-bv+cx-dy, av+bu+cy+dx),$$

$$k(a,b,x,y) = (ka, kb, kx, ky),$$

$$(a,b,x,y)^\circ = (a, -b, x^*, -y^*)$$

for all $a, b, c, d, k\in K$ and $x, y, u, v\in A$.

Proposition 7. *E° is an involution algebra with identity* $(1, 0, 0, 0)$.

$$I^\circ = \{(0, 0, x, y):\ x, y \in A\}$$

is an ideal of E° and $I^\circ \cong A^* \oplus A^{-*}$,

$$L^\circ = \{(0, 0, x, 0):\ x \in A\}$$

is an ideal of I°, $L^\circ \cong A^*$ but L° is not an ideal of E°.

Proof. The proof that E° is an involution algebra, is an exhausting verification and therefore we omit it. The further assertions are straightforward. For the last assertion we notice that by

$$(0, 1, 0, 0)(0, 0, x, 0) = (0, 0, 0, x)$$

we have $E^\circ \cdot L^\circ \nsubseteq L^\circ$.

In the following theorem we shall prove necessary and sufficient conditions for a radical class $\mathbf{R}$ to satisfy A—D—S. This will exhibit the decisive role of the behaviour of involution algebras with zero-multiplication.

Theorem 1. *For a radical class $\mathbf{R}$ of involution algebras the following conditions are equivalent:*

1) *$\mathbf{R}$ satisfies A—D—S,*
2) *if $A^* \in \mathbf{R}$ and $A^2=0$, then $A^\circ \in \mathbf{R}$ for any involution $^\circ$ built on A,*
3) *if $A^* \in \mathbf{R}$ and $A^2=0$, then $A^{-*} \in \mathbf{R}$,*
4) *$A^{\mathrm{id}} \in \mathbf{R}$ if and only if $A^{-\mathrm{id}} \in \mathbf{R}$ whenever $A^2=0$.*

Proof. 1)$\Rightarrow$3) Suppose that condition 3) is not satisfied, that is, there exists an involution algebra A^* such that $A^2=0$, $A^* \in \mathbf{R}$ and $A^{-*} \notin \mathbf{R}$. Obviously the ideals of A^* are exactly those of A^{-*}. Hence without loss of generality we may assume that $A^* \in \mathbf{R}$ and $A^{-*} \in \mathcal{S}\mathbf{R}$. Applying Proposition 7 we get $\mathbf{R}(I^\circ)=L^\circ$ and that L° is not an ideal of E°, though $I^\circ \lhd E^\circ$. Hence $\mathbf{R}$ does not satisfy A—D—S.

Next we show the equivalence of conditions 2), 3) and 4). The implications 2)$\Rightarrow$3)$\Rightarrow$4) are trivial. To prove the implication 4)$\Rightarrow$2), let us assume that $A^* \in \mathbf{R}$, $A^2=0$. First we shall show that $A^{\mathrm{id}} \in \mathbf{R}$. The set

$$D = \{x+x^*:\ x \in A\}$$

is clearly an ideal of A^*, moreover, $D^*=D^{\mathrm{id}}$ holds. The mapping

$$g:\ A^* \to D^*$$

defined by $g(x)=x+x^*$ is obviously a homomorphism of A^* onto D^*, and by $A^* \in \mathbf{R}$ it follows $D^{\mathrm{id}}=D^* \in \mathbf{R}$. In the factor algebra $(A/D)^*$ we have

$$x+D = -x^*+D$$

and

$$(x+D)^* = (-x^*+D)^* = -x+D.$$

Hence we have

$$(A/D)^{-\mathrm{id}} = (A/D)^* = A^*/D^* \in \mathbf{R}.$$

Applying condition 4) it follows $A^{\mathrm{id}}/D^{\mathrm{id}} = (A/D)^{\mathrm{id}} \in \mathbf{R}$. Since $\mathbf{R}$ is closed under extensions, $D^{\mathrm{id}} \in \mathbf{R}$ and $A^{\mathrm{id}}/D^{\mathrm{id}} \in \mathbf{R}$ implies $A^{\mathrm{id}} \in \mathbf{R}$.

Applying condition 4) again, we get $A^{-\mathrm{id}} \in \mathbf{R}$.

Let $^\circ$ denote an arbitrary involution on A and let us consider the set

$$C = \{x+x^\circ : x \in A\}.$$

As above, we get that $C^{-\circ} = C^{-\mathrm{id}}$ is a homomorphic image of $A^{-\mathrm{id}} \in \mathbf{R}$ and hence also $C^{-\mathrm{id}} \in \mathbf{R}$ holds. Since we have also

$$A^\circ/C^{-\mathrm{id}} = A^\circ/C^\circ = (A/C)^{-\mathrm{id}} = A^{-\mathrm{id}}/C^{-\mathrm{id}} \in \mathbf{R},$$

and since $\mathbf{R}$ is closed under extensions, we get $A^\circ \in \mathbf{R}$, proving the validity of condition 2).

Finally we shall show the implication 2)$\Rightarrow$1). Let $I^* \triangleleft A^*$ and $L^* = \mathbf{R}(I^*)$. By 2) and Proposition 4 we have $(L^2/L^3)^* \in \mathbf{R}$ and hence Proposition 5 (iii) yields that $((aL+La^*+L)/L)^*$ is a homomorphic image of the involution algebra $((L/L^2)\oplus(L/L^2))^\square$ which is in $\mathbf{R}$ in view of condition 2) and of $(L/L^2)^* \in \mathbf{R}$. Hence we have $((aL+La^*+L)/L)^* \in \mathbf{R}$. Applying Proposition 5 (ii) it follows that

$$((aL+La^*+L)/L)^* \triangleleft (I/L)^* = I^*/L^* = I^*/\mathbf{R}(I^*) \in \mathcal{S}\mathbf{R}.$$

Thus we get

$$((aL+La^*+L)/L)^* \subseteq \mathbf{R}(I^*/L^*) = 0,$$

that is, $aL+La^* \subseteq L^*$ holds. Hence

$$ax = ax+0a^* \in L \quad \text{and} \quad xa^* = a0+xa^* \in L^*$$

is valid for all $x \in L$. Since the choice of $a \in A$ was arbitrary, we have got $AL \cup \cup LA^* \subseteq L^*$, implying $\mathbf{R}(I^*) = L^* \triangleleft A^*$.

Let us notice that the assertion of Corollary 1 follows immediately also from Theorem 1.

Corollary 2. *For a radical class* $\mathbf{R}$ *the following conditions are equivalent:*

1) $\mathbf{R}$ *satisfies A—D—S,*

5) *if* $A^* \in \mathcal{S}\mathbf{R}$ *and* $A^2=0$, *then* $A^{-*} \in \mathcal{S}\mathbf{R}$,

6) $A^{\mathrm{id}} \in \mathcal{S}\mathbf{R}$ *if and only if* $A^{-\mathrm{id}} \in \mathcal{S}\mathbf{R}$ *whenever* $A^2=0$.

Proof. We show that 3) of Theorem 1 implies 5). Suppose that $A^* \in \mathcal{S}\mathbf{R}$. If $A^{-*} \notin \mathcal{S}\mathbf{R}$, then $0 \neq L^{-*} = \mathbf{R}(A^{-*}) \in \mathbf{R}$. By condition 3) we have $L^* \in \mathbf{R}$. Hence $L^* \subseteq \mathbf{R}(A^*) = 0$ contradicts $L^{-*} \neq 0$.

5)$\Rightarrow$4) Trivial.

6)$\Rightarrow$4) of Theorem 1. This can be proved similarly to the implication 3)$\Rightarrow$5) and therefore it is left to the reader.

Corollary 3. *Let* **R** *be a radical class in* $\mathfrak{V}$. *If all involution algebras of* **R** *with zero-multiplication are of characteristic* 2, *then* **R** *satisfies A—D—S.*

Proof. Condition 3) of Theorem 1 is satisfied, as $x^*=-x^*$ whenever $x\in A^*\in\mathbf{R}$ and $A^2=0$.

For varieties of not-necessarily associative algebras (without involution) over a field satisfying some weaker conditions than the assertions of Propositions 1 and 2, ANDERSON and GARDNER [2] have proved that any radical class **R** either contains all algebras with zero-multiplication or consists of idempotent algebras, and therefore it satisfies A—D—S (cf. [2] Theorem 3. 9). The corresponding assertion for involution algebras is not true, as follows from the following theorem.

Theorem 2. *Let* $\mathfrak{V}$ *be the variety of all involution algebras over a field K. Every radical class in* $\mathfrak{V}$ *satisfies A—D—S if and only if* char $K=2$.

Proof. If char $K=2$, then Corollary 3 yields that every radical class satisfies A—D—S. In the case char $K\neq 2$ on the underlying set K one can build two involution algebras K^{id} and $K^{-\mathrm{id}}$ such that $K^2=0$ and K^{id} is not isomorphic to $K^{-\mathrm{id}}$. As K is a field, K^{id} is a simple involution algebra. Now the upper racical $\mathbf{R}=\mathscr{U}(K^{\mathrm{id}})$ does not satisfy A—D—S because $K^{-\mathrm{id}}\in\mathbf{R}$ and $K^{\mathrm{id}}\in\mathscr{S}\mathbf{R}$.

3. The hereditariness of semisimple classes

If a radical class **R** satisfies A—D—S, then the corresponding semisimple class must be hereditary. The converse of this assertion is not true, and varieties of involution algebras provide natural examples to show that the hereditariness of a semisimple class $\mathscr{S}\mathbf{R}$ does not imply that **R** satisfies A—D—S.

Theorem 3. *A radical class of involution algebras with hereditary semisimple class, need not satisfy A—D—S.*

Proof. We shall construct a radical class **R** which has the desired properties. Let Z denote the algebra of integers (over the ring of integers) with zero-multiplication. The upper radical $\mathbf{R}=\mathscr{U}(Z^{\mathrm{id}})$ does not satisfy A—D—S, as $Z^{-\mathrm{id}}\in\mathbf{R}$ and $Z^{\mathrm{id}}\in\mathscr{S}\mathbf{R}$.

We claim that $A^*=A^{\mathrm{id}}$ for all $A^*\in\mathscr{S}\mathbf{R}$. Suppose that $x^*\neq x$ for some $x\in A$, and consider the ideal B^* of A^* generated by the element x^*-x. Since $A^*\in\mathscr{S}\mathbf{R}$, there exists a homomorphism φ of B^* onto Z^{id}. Hence

$$\varphi(x^*-x) = \varphi(x^*)-\varphi(x) = \varphi(x)^{\mathrm{id}}-\varphi(x) = 0$$

holds, that is, $x^*-x\in\ker\varphi$. This implies $0=B^*/\ker\varphi\cong Z^{\mathrm{id}}$, a contradiction. Thus $A^*=A^{\mathrm{id}}$

If $I^*\lhd A^*\in\mathcal{S}\mathbf{R}$, then by $A^*=A^{\mathrm{id}}$ we have $I^*=I^{\mathrm{id}}$. Hence the standard proof of the hereditariness of semisimple classes of algebras (without involution) works, yielding $I^*\in\mathcal{S}\mathbf{R}$ (cf. [1] or [9]).

We shall see that a semisimple class $\mathcal{S}\mathbf{R}$ need not be hereditary, even if $\mathbf{R}$ is a hereditary radical class. Prior to this we prove some assertions.

Proposition 8. *Let A^* be an involution algebra such that $2a\neq 0$ whenever $0\neq a\in A$. If A^* has an ideal I^* such that $I^*=I^{\mathrm{id}}$ and $(A/I)^*=(A/I)^{\mathrm{id}}$, then also $A^*=A^{\mathrm{id}}$ holds.*

Proof. In $(A/I)^{\mathrm{id}}$ we have

$$x^*+I = (x+I)^* = (x+I)^{\mathrm{id}} = x+I,$$

yielding $x^*-x\in I$ for all $x\in A$. Hence

$$x^*-x = (x^*-x)^{\mathrm{id}} = (x^*-x)^* = x-x^*$$

holds, so we get $2(x^*-x)=0$ for all $x\in A$. By the assumption we conclude that $x^*=x$, that is, $A^*=A^{\mathrm{id}}$.

The Tangeman—Kreiling [7] lower radical construction carries over to involution algebras without difficulty. Given a class $\mathbf{C}$ of involution algebras, define inductively

$\mathbf{C}_1 = \{A^*\colon A^*$ is a homomorphic image of an involution algebra $B^*\in\mathbf{C}\}$,

$\mathbf{C}_\lambda = \{A^*\colon$ there exists an $I^*\lhd A^*$ such that $I^*\in\mathbf{C}_{\lambda-1}$ and $(A/I)^*\in\mathbf{C}_{\lambda-1}\}$

if $\lambda-1$ exists, and

$\mathbf{C}_\lambda = \{A^*\colon A^*$ is the union of an ascending chain of ideals each belonging to one of the classes $\mathbf{C}_\mu$, $\mu<\lambda\}$

if λ is a limit ordinal. Then the smallest radical class containing $\mathbf{C}$, called the *lower radical* of $\mathbf{C}$, is given as

$$\mathcal{L}\mathbf{C} = \cup(C_\lambda \text{ for all ordinals}).$$

If, in addition, $\mathbf{C}$ is a hereditary class of involution algebras, then so is the lower radical $\mathcal{L}\mathbf{C}$.

Proposition 9. *Let $\mathbf{C}$ be the class of involution algebras such that in each $A^*\in\mathbf{C}$, $2a=0$ implies $a=0$. If $A^*=A^{\mathrm{id}}$ for each $A^*\in\mathbf{C}$, then $A^*=A^{\mathrm{id}}$ holds also for every $A^*\in\mathcal{L}\mathbf{C}$.*

Proof. By Proposition 8 each class $\mathbf{C}_\lambda$ consists of algebras with involution id. This yields the assertion.

Corollary 4. *In the variety* $\mathfrak{V}$ *of involution algebras over a field* K *with* char $K\neq 2$, *the class*

$$\mathbf{I} = \{A^*\in \mathfrak{V}\colon\ A^* = A^{\mathrm{id}}\}$$

is a radical class.

Proof. By Proposition 9 we have $\mathbf{I}=\mathscr{L}\mathbf{I}$.

Theorem 4. *Let* $\mathfrak{V}$ *be the variety of all involution algebras over a ring* K *and assume that a homomorphic image* K/M *of* K *is an integral domain such that the quotient field* F *of* K/M *is not of characteristic* 2. *In* $\mathfrak{V}$ *there exists a hereditary radical* $\mathbf{R}$ *whose semisimple class* $\mathscr{S}\mathbf{R}$ *is not hereditary.*

Let us notice that the condition imposed on K is relatively mild as it includes, for instance, the case when K is a subdirect sum of subdirectly irreducible rings such that at least one of them is a field of characteristic $\neq 2$.

Proof. Let us build an involution algebra E° on the set $E=F\times F\times F\times F$ similarly as before Proposition 7 with involution defined by

$$(a, b, x, y)^\circ = (a, -b, x, -y).$$

The assertions of Proposition 7 remain valid. Using the notations of Proposition 7 we have that

$$L^\circ = \{(0, 0, x, 0)\colon\ x\in F\}$$

and so

$$L^{-\circ} = \{(0, 0, 0, y)\colon\ y\in F\}.$$

Further $L^\circ\triangleleft I^\circ\triangleleft E^\circ$ holds, but L° is not an ideal of E°. We claim that I° is a maximal ideal of E°. Let J° be an ideal of E° such that $I^\circ\subsetneqq J^\circ$, and let $(a, b, x, y)\in J^\circ$ be any element of $J^\circ\setminus I^\circ$. We have

$$(a, b, 0, 0) = (a, b, x, y)-(0, 0, x, y)\in J^\circ.$$

If $a\neq 0$, then

$$(a, -b, 0, 0) = (a, b, 0, 0)^\circ\in J^\circ$$

yields $(2a, 0, 0, 0)\in J^\circ$. Hence

$$(1, 0, 0, 0) = (2a, 0, 0, 0)((2a)^{-1}, 0, 0, 0)\in J^\circ$$

holds, implying $J^\circ=E^\circ$. If $a=0$, then $b\neq 0$ and so

$$(1, 0, 0, 0) = (0, b, 0, 0)(0, -b^{-1}, 0, 0)\in J^\circ$$

is valid, implying $J^\circ=E^\circ$. Thus $(E/I)^\circ$ is a simple idempotent involution algebra.

Let $\mathbf{F}$ denote the class consisting of all ideals of the involution algebra L°. The lower radical $\mathbf{R}=\mathscr{L}\mathbf{F}$ of $\mathbf{F}$ is hereditary and by Proposition 9 $A^*=A^{\mathrm{id}}$ holds for every $A^*\in\mathbf{R}$. Moreover, as $(E/I)^\circ$ is simple and idempotent, it follows $(E/I)^\circ\in\mathcal{S}\mathbf{R}$ and so $\mathbf{R}(E^\circ)\subseteqq I^\circ$. As $\mathbf{R}(L^{-\circ})\in\mathbf{R}$, by Proposition 9 we have $x=-x$ for all $x\in\mathbf{R}(L^{-\circ})$. Since char $F\neq 2$, it follows $x=0$, and so $L^{-\circ}\in\mathcal{S}\mathbf{R}$. Hence by the hereditariness of $\mathbf{R}$ we obtain

$$\mathbf{R}(E^\circ)=I^\circ\cap\mathbf{R}(E^\circ)\subseteqq\mathbf{R}(I^\circ)=L^\circ,$$

and the inclusion must be proper as L° is not an ideal in E°. Let $(0,0,x,0)$ be an arbitrary element of $\mathbf{R}(E^\circ)$. If $x\neq 0$, then for each $f\in F$ we have

$$(0,0,f,0)=(fx^{-1},0,0,0)(0,0,x,0)\in\mathbf{R}(E^\circ)$$

implying $\mathbf{R}(E^\circ)=L^\circ$, which is impossible. Hence $\mathbf{R}(E^\circ)=0$ and so $\mathbf{R}(I^\circ)\subsetneqq\mathbf{R}(E^\circ)$ holds. This means that $\mathcal{S}\mathbf{R}$ is not hereditary.

Recall that a class $\mathbf{C}$ of algebras (with or without involution) is said to be a *coradical class,* if $\mathbf{C}$ is hereditary and closed under subdirect sums and extensions. In [2] Anderson and Gardner posed the question whether the concepts "semisimple class" and "coradical class" coincide in a variety of rings satisfying some weaker conditions than the assertions of Proposition 1 and 2. [5] Example 1, 9 of Salavová or our Theorem 4 gives a negative answer to this question.

Corollary 5. *A semisimple class of involution algebras need not be a coradical class.*

References

[1] T. Anderson, N. Divinsky and A. Suliński, Hereditary radicals in associative and alternative rings, *Canad. J. Math.*, **17** (1965), 594—603.

[2] T. Anderson and B. J. Gardner, Semi-simple classes in varieties satisfying an Andrunakievich Lemma, *Bull. Austral. Math. Soc.*, **18** (1978), 187—200.

[3] B. J. Gardner, Extension-closure and attainability for varieties of algebras with involution, *Comment. Math. Univ. Carolinae*, **21** (1980), 285—292.

[4] W. G. Leavitt and R. Wiegandt, Torsion theory for not necessarily associative rings, *Rocky Mountain J. Math.*, **9** (1979), 259—271.

[5] K. Salavová, Radicals of rings with an involution. 1, *Comment. Math. Univ. Carolinae*, **18** (1977), 367—381.

[6] K. Salavová, Radicals of rings with an involution. 2, *Comment. Math. Univ. Carolinae*, **18** (1977), 455—467.

[7] R. Tangeman and D. Kreiling, Lower radicals in nonassociative rings, *J. Austral. Math. Soc.*, **14** (1972), 419—423.

[8] J. Wichmann, Primitive ideals in complex *-algebras, (preprint, 1983).

[9] R. Wiegandt, *Radical and semisimple classes of rings*, Queen's papers in pure and appl. math., 37 (Kingston, Ontario, 1974).

MATHEMATICAL INSTITUTE OF THE
HUNGARIAN ACADEMY OF SCIENCES
REÁLTANODA U. 13—15
1053 BUDAPEST, HUNGARY

Acta Sci. Math., 50 (1986), 15—38

Does a given subfield of characteristic zero imply any restriction to the endomorphism monoids of fields?

PÉTER PRŐHLE

Introduction

E. NOETHER asked whether the Galois groups of normal extensions of the field of rationals can be prescribed. ŠAFAREVIČ showed that each solvable group occures as a Galois group. J. DE GROOT [9] proved that the automorphism groups of rings can be prescribed. More detail: For each group there exists a suitable ring the automorphism group of which is isomorphic to the given group. So J. de Groot asked whether the automorphism groups of fields can be prescribed, too. After a negative result due to Krull, and after a partial solution due to W. KUYK [13] the question was answered by E. FRIED—J. KOLLÁR [5]. As a corollary of a much stronger result it was shown that: To each group G there exists a field F of a given characteristic different from 2, where G is isomorphic to the automorphism group of F. Each field given by the construction in [5] is a transcendental extension of its own prime field. It came also to light that the procedure used in [5] is unfit for handling the extensions of algebraically closed fields. So has been raised the question formulated in the title above. The answer to the analogous question is affirmative with respect to the class of graphs by L. BABAI—J. NEŠETŘIL [2], to the class of bounded lattices by M. E. ADAMS—J. SICHLER [1], to the class of unary algebras by J. KOLLÁR [11, 12] and to the class of integral domains of characteristic zero by E. FRIED [4]. This paper presents a solution in the case of fields of characteristic zero and in the case of non-unary algebras.

Received February 6, 1984.

The results

If the only endomorphism of a structure is the identity, then the structure is called rigid. A monoid is called right cancellative, if $xz=yz$ implies $x=y$.

Theorem 1. *Each field of characteristic zero is embeddable into a rigid one.*

Theorem 2. *Let F be a given field of characteristic zero. Then a monoid M is isomorphic to the endomorphism monoid of a field containing F as a subfield iff M is right cancellative.*

A functor $F\colon A\to B$ which is injective on every $\mathrm{Hom}_A(a, a'')$ is called faithful. If, in addition, F is also injective on the class of all objects of A we call it an embedding. F is full if every morphism $d\colon F(a)\to F(a'')$ of B has the form $d=F(c)$ for some morphism $c\colon a\to a''$ from A. A concrete category is a category A together with a fixed faithful functor $U\colon A\to \mathrm{SET}$, where SET is the category of all sets and all mappings. A category of structures will always be considered as a concrete category whose faithful functor is the usual "underlying set" functor.

Let Fields, $\mathrm{Alg}(t)$ and $\mathrm{Rel}(t)$ denote the category whose objects are the fields of characteristic zero, the algebras of the given similarity type t and the relational structures of type t, and whose morphisms are the 1-preserving ring homomorphisms, the usual homomorphisms and the weak homomorphisms. Let C be a concrete category, then $\mathrm{Inj}\, C$ denotes the subcategory of those morphisms of C, which are carried by injective mappings. For $a\in \mathrm{Ob}(C)$, $\mathrm{Ext}(a, C)$ denotes the full subcategory of those objects of C, which have a as a subobject.

Let A and B be concrete categories and let $U\colon A\to \mathrm{SET}$ and $V\colon B\to \mathrm{SET}$ be their corresponding "underlying set" functors. A full embedding $F\colon A\to B$ is called an extension if there is a monotransformation from U into $V\circ F$. F is a strong embedding if $H\circ U=V\circ F$ for some faithful functor $H\colon \mathrm{SET}\to \mathrm{SET}$, here H is called the carrier of F. It is easy to see that a functor $H\colon \mathrm{SET}\to\mathrm{SET}$ is faithful iff there is a monotransformation from the identity functor on SET into the functor H. Thus every strong embedding is also an extension.

Theorem 3. *Let F be a given field of characteristic zero. Then* $\mathrm{Inj}\,\mathrm{Alg}(t)$ *has a strong embedding into* $\mathrm{Ext}(F, \mathrm{Fields})$ *and* $\mathrm{Inj}\,\mathrm{Rel}(t)$ *has an extension into* $\mathrm{Ext}(F, \mathrm{Fields})$, *for each similarity type t.*

Theorem 4. *Let A be an algebra of similarity type t, where t contains at least one at least binary operation. Then the following statements are equivalent:*

(a) *A has no one-element subalgebra;*

(b) *A is embeddable into a rigid algebra of similarity type t;*

(c) $\mathrm{Alg}(s)$ *has a strong embedding into* $\mathrm{Ext}(A, \mathrm{Alg}(t))$ *and* $\mathrm{Rel}(s)$ *has an extension into* $\mathrm{Ext}(A, \mathrm{Alg}(t))$, *for each similarity type s.*

Review of the technique

For the basic notions and for the customary technique see the textbook of G. Grätzer [8], of S. MacLane [14], of A. Pultr—V. Trnková [15] and of B. L. van der Waerden [19], and the paper of P. Vopenka—A. Pultr—Z. Hedrlin [18]. For the technique of rings and fields see E. Fried [3, 4], E. Fried—J. Kollár [5], E. Fried—J. Sichler [6, 7] and J. Kollár [10].

To prove Theorem 1 we want to mark the elements of the given field by special extensions. Namely, two elements can be transposed by an automorphism only if they have isomorphic marks. If we want to mark a subset A of an integral domain I, then it is enough to use the following process of extension due to E. Fried [3]: First take the algebraic closure of I. Then take the polynomial ring in one variable over this algebraic closure. Finally add the reciprocial of the polynomials of the form $(y-a)$ where a runs over the given subset A. It can be shown that each automorphism fixes the variable y, and the set A is permuted only. Of course, this statement holds only for some carefully chosen sets A, see [3]. In the case of fields we must take the whole quotient field of the polynomial ring. But we may try to add the square roots of the polynomials in question (the square roots of the polynomials of the form $(y-a)$. It is easy to see that this modification is insufficient: the variable y can be moved by the endomorphisms. On the other hand there are a lot of flip-flops: namely the conjugates of the roots can be permuted. To prevent the motion of the variable y we take an odd prime p and we make the element y p-high by adding its p-th roots. So we get a bigger field. If the unit element is the only p-high element of the original smaller field, then this extension really denotes the set A. But it is easy to show that if the smaller field contains an algebraically closed field — this is the general case — then this extension doesn't mark the set A. Therefore we add not only the square roots of the polynomials $(y-a)$, but also the square roots of $(y-1)$, $(y-a^{11})$.

If we want to embed an uncountable algebraically closed field into a rigid field (see Theorem 1), then this rigid field must have a greater cardinality than the original algebraically closed field. We also see that the construction above doesn't change the cardinality of the fields, even if we iterate that process for other odd primes. For the simple reason that we may enlarge the cardinalities the final form of the extension we will use in the proofs is just the special extension $F(E, Y, p, q)/F$, the definition of which can be found in the first part of the main lemma.

The investigation of the special extension

Main lemma (first part). *Let F be a field of characteristic zero, Y be a set disjoint to F, E be a subset of $F\times Y$, and p and q be two different primes. We use the following notations for $y\in Y$: ${}_0y=y$,*

$$A(y)=\{a\colon \langle a,y\rangle\in E\} \quad \text{and} \quad B(y)=\{1,a,a^{11}\colon \langle a,y\rangle\in E\}.$$

Then the following property uniquely determines a field denoted by $F(E,Y,p,q)$: $F(E,Y,p,q)$ is the extension of F generated by the set

$$R=\{{}_iy,\ t(b,y)\colon i\in\omega,\ b\in B(y);\ y\in Y\},$$

where:

(a) *Y is an algebraically independent system over F,*

(b) *$({}_{i+1}y)^p={}_iy$, for $i\in\omega$ and $y\in Y$,*

(c) *$(t(b,y))^q=y-b$, for $b\in B(y)$ and $y\in Y$.*

We shall use the following occasional symbolic nomenclature:

$F(E,Y,p,q)$	special extension
$F(E,Y,p,q)\setminus F$	the skin of the extension
Y	the variables of the skin
(F,E,Y)	the bipartite graph of the skin
R	the roots of the skin.

Let $F(E,Y,p,q)$ and $F''(E'',Y'',p,q)$ be two special extensions. A mapping $f\colon F\cup Y\to F''\cup Y''$ is said to be a pre-morphism if f is injective, $f|_F$ is an embedding of F into F'', and f is a homomorphism of the bipartite graph (F,E,Y) into (F'',E'',Y''). A mapping $f\colon F\cup R\to F''\cup R''$ is said to be a pre-homomorphism if $f|_{F\cup Y}$ is a pre-morphism, $f({}_iy)={}_i(f(y))$ and $f(t(b,y))=t(f(b),f(y))$ for $i\in\omega$, $b\in B(y)$ and $y\in Y$. A field homomorphism of $F(E,Y,p,q)$ into $F''(E'',Y'',p,q)$ sending the subfield F into the subfield F'' and the set R into R'' is called a special homomorphism. If the two special extensions in question are the same, then we can use the expression "endo" instead of "homo".

Main lemma (second part). *Let us take two special extensions: $F(E,Y,p,q)$ and $F''(E'',Y'',p,q)$, where each of the sets $A(y)$ and $A''(y'')$ is an algebraically independent system over the prime field, for $y\in Y$ and $y''\in Y''$ respectively. Then*

(a) *For each special homomorphism h of $F(E,Y,p,q)$ into $F''(E'',Y'',p,q)$ the restriction of h to $F\cup Y$ is a pre-morphism, and the restriction of h to $F\cup R$ is a pre-homomorphism.*

(b) *Each pre-morphism has a unique extension which is a special homomorphism.*

(c) *The category whose objects are the special extensions and whose morphisms are the special homomorphisms is naturally equivalent to the category whose objects are the special extensions and whose morphisms are the pre-morphisms.*

Abel's theorem. *A polynomial* (x^k-b) *of prime degree* k *over a field* L *is reducible if, and only if,* b *is a* k^{th}*-power in* L.

A simple proof can be found in the texbook of L. Rédei [16].

Lemma 1. *Let* L *be a field of characteristic zero. Take the simple algebraic extension* $L(t)$, *where* (x^n-t^n) *is an irreducible polynomial in the polynomial ring* $L[x]$, *and* n *is a prime. Let* m *be an integer greater than* 1. *Then the* m^{th}*-power of an element of* $L(t)$ *belongs to the subfield* L *iff the element is of the form* $c\cdot t^k$, *where* $c\in L$, $0\leqq k<n$ *and* $n|km$. *If in addition* $(m,n)=1$ *then an element of* L *has an* m^{th}*-root in* $L(t)$ *iff it has one in* L.

Proof. Let K be the smallest algebraic extension of L containing all the n^{th}-roots of unity. The degree of the extension K/L is less than n. So the irreducibility of (x^n-t^n) over L implies that t^n is never an n^{th}-power in K. Consequently by Abel's theorem (x^n-t^n) is irreducible over K. So any element b of $K(t)$ can be uniquely written in the form $b_0+b_1t+b_2t^2+\cdots\cdots+b_{n-1}\cdot t^{n-1}$, where all the coefficients belong to K. Obviously $L(t)\subseteqq K(t)$, and an element b of $K(t)$ belongs to $L(t)$ iff each of the coefficients of b belongs to L. Let u be a primitive n^{th}-root of unity. The mapping $t\mapsto u\cdot t$ induces a relative automorphism of the extension $K(t)/K$, where the image of b is: $b_0+b_1\cdot u\cdot t+b_2\cdot u^2\cdot t^2+\dots\dots+b_{n-1}\cdot u^{n-1}\cdot t^{n-1}$.

If $b^m\in L$, then this image of b must be $v\cdot b$, where v is a suitable m^{th}-root of unity. The uniqueness of the coefficients of $v\cdot b$ gives the following equations: $b_i(u^i-v)=0$ for $0\leqq i<n$. If $b\neq 0$, then there is an index k for which $b_k\neq 0$. Consequently $u^k-v=0$, and $b_i=0$ for $i\neq k$. So $b=b_k\cdot t^k$, where $b\in L(t)$ implies $b_k\in L$. Further $n|km$, since $u^{km}=v^m=1$. $(m,n)=1$ yields $k=0$.

Lemma 2. *Let* K *be a transcendental extension of* L *such that* K *is an algebraic extension of finite degree with respect to the simple transcendental extension* $L(y)$. *Let* s *be a prime. An element is called s-high in a field, if the element has an* s^j*-th root in the field for each* $j\in\omega$. *Then each s-high element of* K *belongs to* L.

Proof. Let $x\in K\setminus L$. Then y is algebraic over $L(x)$, so K is an algebraic extension of finite degree with respect to $L(x)$. Suppose, that x is s-high. Let ${}_ix$ be an s^i-th root of x. Consider the infinite chain $L(x)\leqq L({}_1x)\leqq L({}_2x)\leqq\dots\leqq L({}_ix)\leqq\dots$ of fields. As the degree of $K/L(x)$ is finite, there exists an index n such that $L({}_nx)=$ $=L({}_{n+1}x)$. So the transcendental element ${}_nx$ has an s^{th}-root in $L({}_nx)$, but that is impossible. So any s-high element must belong to L.

Lemma 3. *Let* F *be a field of characteristic zero. Let* p *and* q *be two different primes. Suppose that* K *is an extension of* F *generated by the set* $\{{}_iz,\ t_v\colon i\in\omega,\ v\in V\}$, *where:*

(a) ${}_0z$ *is transcendental over* F,

(b) $({}_{i+1}z)^p = {}_iz$, $i \in \omega$,

(c) *the elements* $T_v = (t_v)^q$ *are polynomials from the polynomial ring* $F[{}_0z]$, *such that they are mutually prime, and none of them is a constant, nor is divisible by* ${}_0z$, *nor has multiple factor.*

Denote the subfield $F(\{{}_iz, t_v\colon i<j,\ v\in W\})$ *of* K *by* $F(j, W)$, *for* $W \subseteqq V$ *and* $j=1, 2, \ldots, n, \cdots, \omega$. *Then the field* K *has the following properties:*

(1) *The polynomial* $(x^p - {}_iz)$ *is irreducible over* $F(i+1, W)$, *for* $W \subseteqq V$ *and* $i \in \omega$.

(2) *The polynomial* $(x^q - T_v)$ *is irreducible over* $F(j, W)$, *for* $W \subseteqq V$, $v \in V \setminus W$ *and* $1 \leqq j \leqq \omega$.

(3) *If the* q^{th}*-power of an element of* $F(j, W)$ *belongs to the subfield* $F(k, \emptyset)$, *where* $W \subseteqq V$ *and* $k \leqq j \leqq \omega$, *then the element can be written in the form*

$$c\big(f({}_iz)/g({}_iz)\big) \prod_{w\in W'} (t_w)^{n_w},$$

where $c \in F$, f *and* g *are mutually prime polynomials over* F *and both of them have leading coefficients* 1, $i \leqq k$, W' *is a suitable finite subset of* W, *and* $0 < n_w < q$ *for* $w \in W'$.

(4) K *is a transcendental extension of* F.

(5) *Each s-high element of* K *belongs to* F *whenever* s *is a prime different from* p *and* q.

(9) *Each p-high element of* K *is of the form* $c \cdot ({}_iz)^m$, *where* c *is a p-high element of* F, $i \in \omega$ *and* m *is an integer.*

PROOF. Proposition 1 of E. FRIED [3] and Propositions 16, 23 and 24 of E. FRIED—J. KOLLÁR [5] essentially cover the case $q=2$ of the above lemma.

First step: we prove the properties (2) and (3) in the case of finite W and $j=k=1$. We prove by induction on the size of the set W.

$F(1, \emptyset)$ is the quotient field of the polynomial ring $F[{}_0z]$, therefore the property (3) is true in the case of $W=\emptyset$ and $j=k=1$. If the property (3) is true for W and $j=k=1$, then $t_v \in F(1, W)$ would imply an equality of the form

$$g^q({}_0z) \cdot T_v = c^q \cdot f^q({}_0z) \cdot \prod_{w\in W} (T_w)^{n_w}, \quad \text{if} \quad v \in V \setminus W.$$

However, this contradicts one of the conditions on the polynomials T_w. Hence, $t_v \notin F(1, W)$ yields the property (2), by Abel's theorem, for the case of the same W and $j=1$. Now suppose, that both of the properties (2) and (3) are true for a finite W and $j=k=1$. Let $v \in V \setminus W$, $b \in F(1, W \cup \{v\})$ and $b^q \in F(1, \emptyset)$. As $(x^q - T_v)$ is irreducible over $F(1, W)$ by the assumption, $b = c \cdot (t_v)^n$ by the Lemma 1. Here $c \in F(1, W)$ and $c^q \in F(1, \emptyset)$, so the form of c is known by the assumption. Consequently, b has the desired form, too. So we get the property (3) for the index set $W \cup \{v\}$ and $j=k=1$.

Second step: we prove the property (1) in the case of finite W and $i=0$, by induction on the size of the set W.

By Abel's theorem it is enough to show that ${}_0z$ has no p^{th}-root in $F(1; W)$. The existence of a p^{th}-root of ${}_0z$ in $F(1, \emptyset)$ would imply a polynomial equation $g^p({}_0z)\cdot {}_0z=c^p\cdot f^p({}_0z)$, where f and g are mutually prime, which is a contradiction. Lemma 1 gives the inductional step of the proof, as we have seen the irreducibility of (x^q-T_v) over $F(1, W)$ for finite W.

Third step: we prove the properties (1) and (2).

If we replace the elements ${}_0z; {}_1z, {}_2z, \dots$ with ${}_iz, {}_{i+1}z, {}_{i+2}z, \dots;$ then the conditions in Lemma 3 remain satisfied. So the polynomials (x^q-T_v) and $(x^p-{}_iz)$ are irreducible over $F(i+1, W)$ for finite $W\subseteq V$, $v\in V\setminus W$ and $i\in\omega$. The reducibility of a polynomial over a field L needs only a finitely many coefficiences from L, therefore a reducible polynomial is also reducible over a suitable finitely generated subfield of L. So we get the properties (1) and (2) by an indirect proof.

Fourth step: we prove the property (3).

As the polynomial $(x^p-{}_iz)$ is irreducible over $F(i+1, W)$ for $i\in\omega$, Lemma 1 shows that if the q^{th}-power of an element of $F(i+2, W)$ belongs to $F(i+1, W)$; then the element also belongs to $F(i+1, W)$. So, if an element of $F(1, \emptyset)$ has a q^{th}-root in $F(\omega, W)$, then this q^{th}-root belongs to $F(1; W)$. However, ${}_iz$ can get the rôle of ${}_0z$. Consequently we get the property (3) for finite j, k and W. Finally each element of $F(j, W)$ belongs to a field $F(i+1, W')$ for suitable finite $W'\subseteq W$ and $i<j$.

Fifth step: we prove the property (4).

Let x be an algebraic element of K over F. Let $L=F(x)$. The element ${}_0z$ is transcendental over L, since x is algebraic. All the other conditions of Lemma 3 are also satisfied with respect to L instead of F. Therefore, the system $1, t_v, (t_v)^2, \dots \dots, (t_v)^{q-1}$ forms a basis of the field extension $\hat{L}=L(\omega, W\cup\{v\})/L(\omega, W)$, for $v\in V\setminus W$, satisfying the following property: an element of $\hat{L}$ belongs to $F(\omega; W\cup\{v\})$ iff the coefficients of the element with respect to this basis belong to $F(\omega, W)$. Consequently, $x\in F(\omega, W\cup\{v\})$ implies $x\in F(\omega, W)$, since the coefficients of x must belong to $F(\omega, W)$ and $x\in L<L(\omega; W)$. So $x\in F(\omega, \emptyset)$. Therefore $x\in F({}_iz)$ for a suitable $i\in\omega$. But $F({}_iz)$ is a pure transcendental extension of F, so $x\in F$.

Sixth step: we prove the property (5).

Let x be s-high in K. Clearly, $x\in F(i, W)$ for a suitable $i\in\omega$ and a finite $W\subseteq V$. Using Lemma 1 and properties (1) and (2) we get that x is s-high in $F(i, W)$, too. Now we can apply Lemma 2 for $F(i, W)$, so $x\in F$.

Seventh step: we prove the property (6).

Let x be a p-high element of K. Then $x\in F(\omega, W)$ for a suitable finite $W\subseteq V$. By Lemma 1 and by the property (2) x is p-high in the subfield $F(\omega, W)$, too.

So it is enough to prove the following statement by induction on the size of the set W: For finite $W \subseteq V$ the p-high elements of $F(\omega, W)$ are of the form $c \cdot ({}_iz)^m$.

If $W = \emptyset$, then $x \in F(i+1, \emptyset) = F({}_iz)$ for suitable $i \in \omega$. So $x = ({}_iz)^m \cdot (f({}_iz)/g({}_iz))$, where m is an integer, ${}_iz \nmid f({}_iz)$ and ${}_iz \nmid g({}_iz)$. Here $(f({}_iz)/g({}_iz))$ must be p-high in $F(\omega, \emptyset)$. Suppose that there exists an element $y \in F(\omega, \emptyset) \setminus F(i+1, \emptyset)$ such that $y^p \in F(i+1, \emptyset)$ and some p^j-th power of y is $(f({}_iz)/g({}_iz))$. Let $k = \max\{n : y \notin F(n+1, \emptyset)\}$. By Lemma 1 and by the property (1) $y = ({}_{k+1}z)^b \cdot (u({}_kz)/v({}_kz))$, where u and v are polynomials over F, and $0 < b < p$. Now, we arrive at the equation

$$({}_kz)^{b \cdot p^{j-1}} \cdot (u({}_kz))^{p^j} \cdot g({}_iz) = f({}_iz) \cdot (v({}_kz))^{p^j}$$

in the polynomial ring $F[{}_kz]$. Consider the powers of the irreducible factor ${}_kz$ in that equation. As ${}_iz$ is irreducible in $F[{}_iz]$, therefore ${}_iz \nmid f({}_iz)$ implies $({}_iz, f({}_iz)) = 1$ in $F[{}_iz]$. So $({}_iz, f({}_iz)) = 1$ in $F[{}_kz]$, too. Consequently ${}_kz \nmid f({}_iz)$, and by a similar argument ${}_kz \nmid g({}_iz)$. The exponent of ${}_kz$ in $({}_kz)^{b \cdot p^{j-1}} \cdot (u({}_kz))^{p^j} \cdot g({}_iz)$ is congruent to $b \cdot p^{j-1}$ modulo p^j, while the exponent of ${}_kz$ in $f({}_iz) \cdot (v({}_kz))^{p^j}$ is divisible by p^j. This is a contradiction, and so, in opposit our assumption, the quotient $(f({}_iz)/g({}_iz))$ must be p-high even in $F(i+1, \emptyset)$. Therefore by Lemma 2 $(f({}_iz)/g({}_iz)) \in F$, consequently x has the form $c \cdot ({}_iz)^m$, what we had to prove.

Now we suppose that there exists a $w \in W$, and the statement is true for $W \setminus \{w\}$. Let $L = F(\omega, W \setminus \{w\})$ and $K = F(\omega, W)$. By the property (2) the degree of the extension K/L is q. Let $N(d)$ denote the norm of d with respect to K/L for $d \in K$. Only the following property of the norm will be used: N is a multiplicative mapping from K into L such that $N(d) = d^q$ for $d \in L$. For the details see L. Rédei [16] and B. L. van der Waerden [19]. $N(x)$ is p-high in L, as x is p-high in K. So the element $y = x^q/N(x)$ is p-high in K. Clearly $y \in F(i+1, W)$ for a suitable $i \in \omega$. Suppose that there exists an element $u \in F(\omega, W) \setminus F(i+1, W)$ such that $u^p \in F(i+1, W)$ and y is a p^j-th power of u. Let $k = \max\{n : u \notin F(n+1, W)\}$. By Lemma 1 and by property (1) $u = h \cdot ({}_{k+1}z)^b$, where $h \in F(k+1, W)$ and $0 < b < p$. $N(u) = N(h) \cdot (N({}_{k+1}z))^b = N(h) \cdot ({}_{k+1}z)^{b \cdot q}$. So $N(u) \notin F(k+1, W)$, as $N(h) \in F(k+1, W)$ and $p \nmid bq$. However, $N(y)$ is the p^j-th power of $N(u)$, and $N(y) = N(x^q/N(x)) = (N(x))^q/N(N(x)) = 1$. This is a contradiction, because by the property (4) $N(u) \notin F(k+1, W)$ implies that $N(u)$ is a transcendental element, while its p^j-th power should be 1. Therefore, in opposit our assumption, y must be p-high even in $F(i+1, W)$. Consequently, by Lemma 2 $y \in F$, and therefore $y \cdot N(x)$ is a p-high element of L. So by the inductional hypothesis $y \cdot N(x)$ has the form $c \cdot ({}_iz)^m$. Using the property (3), we get:

$$c \cdot ({}_iz)^m = y \cdot N(x) = x^q = d^q \cdot (f^q({}_iz)/g^q({}_iz)) \prod_{w \in W} (T_w)^{n_w}.$$

This implies, that $n_w = 0$, $g({}_iz) = 1$, $q \mid m$ and $f({}_iz) = ({}_iz)^{(m/q)}$. So x also has the desired form: $x = d \cdot ({}_iz)^{(m/q)}$.

Proof of the first part of the main lemma. First of all we fix a well ordering $(Y; <)$ of the variables. For $y \in Y$ let

$$F_y = F(\{{}_iu,\ t(b,u)\colon\ i\in\omega,\ b\in B(u),\ u\in Y \text{ and } u<y\})$$

and

$$K_y = F_y(\{{}_iy,\ t(b,y)\colon\ i\in\omega,\ b\in B(y)\}).$$

The special extension $F(E, Y, p, q)$ must be the union of the ascending chain of the subfields K_y, so it is enough to prove the unique existence of the subfields F_y and K_y by transfinite induction on $y\in(Y, <)$.

If $y\in Y$ is the least element of $(Y, <)$; then F_y must be F. If y is not the least element of $(Y, <)$; then F_y must be $\bigcup\{K_u\colon u\in Y,\ u<y\}$, where the subfields K_u form an ascending chain. Finally we show that K_y uniquely exists; whenever F_y does.

Now, y is transcendental over F_y, because y is transcendental over $F(\{u\colon u\in Y,\ u<y\})$ and F_y is an algebraic extension of $F(\{u\colon u\in Y,\ u<y\})$. By the conditions for i, $({}_{i+1}y)^p={}_iy$. The elements $y-b=(t(b,y))^q$ are polynomials from the polynomial ring $F_y[y]$ for $b\in B(y)$, where none of them is a constant, none of them is divisible by y, none of them has a multiple factor and they are mutually prime. So Lemma 3 can be used for the extension K_y of F_y. By the property (1) $F_y({}_{i+1}y)$ must be the simple algebraic extension of $F_y({}_iy)$ by the root of the irreducible polynomial $(x^p-{}_iy)$ for $i\in\omega$. Further $F_y(\{{}_iy\colon i\in\omega\})$ must be the union of the ascending chain

$$F_y \leqq F_y(y) \leqq F_y({}_1y) \leqq F_y({}_2y) \leqq \ldots \ldots \leqq F_y({}_iy) \leqq \ldots.$$

Now we fix a well ordering $(B(y), <)$. Let $F_{\substack{\leqq\\ yb}} = F_{\substack{<\\ yb}}(t(b,y))$, where $F_{\substack{<\\ yb}} = F_y(\{t(c,y)\colon c\in B(y),\ c<b\})$ for $b\in B(y)$. Clearly K_y must be the union of the ascending chain of the subfields $F_{\substack{\leqq\\ yb}}$, so it is enough to prove the unique existence of the subfields $F_{\substack{<\\ yb}}$ and $F_{\substack{\leqq\\ yb}}$ by transfinite induction on $b\in(B(y), <)$. If $b\in B(y)$ is the least element of $(B(y), <)$, then $F_{\substack{<\\ yb}}$ must be F_y. If b is not the least element of $(B(y), <)$, then $F_{\substack{<\\ yb}}$ must be $\bigcup\{F_{\substack{\leqq\\ yc}}\colon c\in B(y),\ c<b\}$; where the subfields $F_{\substack{\leqq\\ yc}}$ form an ascending chain. Finally by the property (2) $F_{\substack{\leqq\\ yb}}$ must be the simple algebraic extension of $F_{\substack{<\\ yb}}$ by the root of the irreducible polynomial $(x^q-(y-b))$.

To prove the second part of the main lemma, we need the following four sublemmas. The first three sublemmas have a common condition: Let us take a special extension $F(E, Y, p, q)$, where each set $A(y)$ is an algebraically independent system of elements over the prime field, for $y\in Y$.

Sublemma 1. *Let $Q(x)$ denote the following sentence: There exists a non-zero element u in F and an element w in $F(E, Y, p, q)\setminus F$, where w is p-high in $F(E, Y, p, q)$, $(w-u)$ is the q^{th}-power of an element v of $F(E, Y, p, q)$, and $x=u/w$. Then $Q(x)$*

is equivalent to the following: The bipartite graph of the skin has an edge $\langle a, y\rangle$ *such that* $x\in\{(1/y), (a/y), (a^{11}/y)\}$.

Proof. First of all we fix a well ordering $(Y; <)$ of the variables. Now we use the same notation as in the proof of the first part of the main lemma. Suppose that x is an element satisfying $Q(x)$. Set $y=\min\{z\colon w\in K_z\}$. Lemma 3 will be used for the special extension K_y/F_y. As K_y is algebraically closed with respect to $F(E; Y, p, q)$; w is p-high in K_y and $(w-u)\in K_y$ yields $v\in K_y$. So $w=e\cdot({}_iy)^k$, where e is a non-zero p-high element of F_y, $i\in\omega$ and k is a non-zero integer. It can be supposed, that $p|k$ occurs only if $i=0$. Further

$$v = c\cdot(G({}_iy)/H({}_iy))\cdot\sqrt[q]{({}_0y-b_1)^{k_1}\cdot({}_0y-b_2)^{k_2}\cdot\ldots\cdot({}_0y-b_n)^{k_n}}$$

where $0\neq c\in F_y$, G and H are mutually prime polynomials over F_y both of which have leading coefficients 1, $n\in\omega$, $b_1, b_2, \ldots, b_n$ are different elements from $B(y)$; and $0<k_j<q$ for $j=1, 2, \ldots, n$. Set $t={}_iy$. According to the sign of k we get one of the following equations in the polynomial ring $F_y[t]$:

$$H^q(t)\cdot(e\cdot t^k-u) = c^q\cdot G^q(t)\cdot(t^{p^i}-b_1)\cdot\ldots\cdot(t^{p^i}-b_n) \quad \text{if} \quad k>0$$

$$H^q(t)\cdot(e-(t^{-k})\cdot u) = c^q\cdot G^q(t)\cdot(t^{-k})\cdot(t^{p^i}-b_1)\cdot\ldots\cdot(t^{p^i}-b_n) \quad \text{if} \quad k<0.$$

By the assumption none of the elements $e, u, b_1, b_2, \ldots, b_n$ is zero. Therefore each of the binomials occurring in the equations is a proper binomial, consequently none of them has multiple factor. In both cases $G^q(t)$ divides the binomial standing on the left side, so $G^q(t)=1$. In the first case a similar argument shows that $H^q(t)=1$. In the second case we get only that $H^q(t)|t^{-k}$. But $e\neq 1$ yields that $t^{-k}|H^q(t)$, so $H^q(t)=t^{-k}$. Consequently $q|k$ if $k<0$. Now, in both cases the degree of the left side is $|k|$, and the degree of the right side is $n\cdot p^i$. So $n\neq 0$ and $i=0$, since $k\neq 0$ and $i\neq 0$ would imply $p\nmid k$. Now $n=1$, since the quotient of any different elements of $B(y)$ is never an n^{th}-root of unity. The second case is impossible as $q|k=-n=-1$. So the only possible case is the following: $e\cdot y-u=c^q\cdot(y-b)$. Consequently, we have that $x=u/(e\cdot y)=b_1/y$. The other direction of the equivalence is trivial.

Sublemma 2. *Let* $E(a, y)$ *denote the following sentence: The two elements* a *and* y *are transcendental over the prime field,* $Q(1/y)$, $Q(a/y)$ *and* $Q(a^{11}/y)$ *(the notation is in Sublemma 1). Then* $E(a, y)$ *is equivalent to the following:* $\langle a, y\rangle$ *is an edge of the bipartite graph of the skin.*

Proof. Let the elements a and y satisfy $E(a, y)$. Then, by Sublemma 1 there are variables y_k and elements $b_k\in B(y_k)$ such that $a^k/y=b_k/y_k$ for $k=0, 1, 11$. The equation $(a/y)^{11}=(1/y)^{10}\cdot(a^{11}/y)$ implies that:

$$b_1^{11}/y_1^{11} = (b_0^{10}/y_0^{10})(b_{11}/y_{11}).$$

As the elements b_k are different from zero, each of these three variables y_k is algebraically dependent of the other two over F. So, by the structure of the variables we get that $y_0=y_1=y_{11}$, and therefore $b_1^{11}=b_0^{10}\cdot b_{11}$. Here the algebraic independence of $A(y)$ implies the existence of a suitable $c\in A(y)$ such that $\{b_0, b_1, b_{11}\}\subseteq \subseteq\{1, c, c^{11}\}$. Further b_0, b_1 and b_{11} are different elements, because the three quotients $(1/y)$, (a/y) and (a^{11}/y) are also different. Consequently $\langle b_0, b_1, b_{11}\rangle$ is a permutation of $\langle 1, c, c^{11}\rangle$. So, we have to solve the equation $11i=10j+k$ where $\langle i, j, k\rangle$ is a permutation of $\langle 0, 1, 11\rangle$. The only solution is: $i=1$, $j=0$, $k=11$. So, we arrive at the equations $1/y=1/y_1$, $a/y=c/y_1$ and $a^{11}/y=c^{11}/y_1$. Consequently $y=y_1$ is a variable, and $a=c\in B(y_1)=B(y)$. The other direction of the equivalence is trivially true.

Sublemma 3. *Let $V(y)$ denote the following sentence: $y\neq 0$ and $Q(1/y)$ hold, and for all a and z from $F(E, Y, p, q)$ $E(a, z)$ implies that the both of (a/z) and (a^{11}/z) are different from $(1/y)$. Then $V(y)$ is equivalent to the following: y is a variable of the skin.*

Proof. Let y be an element satisfying $V(y)$. By Sublemma 1 $1/y=b/u$, where u is a suitable variable of the skin and $b\in B(u)$. If $A(u)=\emptyset$, then $B(u)=\{1\}$, and then $b=1$. If $A(u)\neq\emptyset$, then for $a\in A(u)$, $E(a, u)$ and $E(a^{11}, u)$ hold, and therefore both of (a/y) and (a^{11}/y) are different from (b/y). So (even in the case of $A(u)\neq\emptyset$), the only possibility is $b=1$. Consequently, in both cases $y=u$ is a variable. The other direction of the equivalence is trivially true.

Sublemma 4. *Under the condition of the second part of the main lemma suppose that a given homomorphism h of $F(E, Y, p, q)$ into $F''(E'', Y'', p, q)$ maps the subfield F into F''. Let $Q''(x'')$, $E''(a'', y'')$, $V''(y'')$ and $A''(y'')$ be defined similarly for $F''(E'', Y'', p, q)$ as $Q(x)$, $E(a, y)$, $V(y)$ and $A(y)$ are for $F(E, Y, p, q)$. Then the following implications hold:*

(a) *If $h(x)\notin F''$ and $Q(x)$ holds, then $Q''(h(x))$.*

(b) *If $h(y)\notin F''$ and $E(a, y)$ holds, then $E''(h(a), h(y))$.*

(c) *If $h(y)\notin F''$ and $V(y)$ holds, then $Q''(1/h(y))$.*

(d) *If $h(y)\notin F''$ and $V(y)$ holds, then $V''(h(y))$, whenever none of the sets $A(y)$ and $A''(y'')$ is empty.*

Note: in particular, each of these implications holds if h is a special homomorphism.

Proof. (a) The validity of $Q(x)$ is demonstrated by suitable elements u, v and w. The image of these elements demonstrate the validity of $Q''(h(x))$, since $h(w)\notin F''$ by the assumption $h(x)\notin F''$.

(b) We have only to use the definition of $E(a, y)$ and the implication (a) of the present sublemma.

(c) We can use the implication (a), since $V(y)$ implies $Q(1/y)$.

(d) If none of the sets $A(y)$ is empty, then $V(y)$ is equivalent to the formula $\exists a(E(a, y))$. Using this equivalence and the implication (b) we get the implication (d).

Proof of the second part of the main lemma. (a) Let y be an arbitrary variable from Y. Using Sublemma 1 and the implication (c) of Sublemma 4 we get $h(y)=x/b$, where $x \in Y''$ and $b \in B(x)$. But $h(y) \in R''$ implies that $b=1$. Consequently, each special homomorphism maps the set Y into Y''. Clearly the restriction $h|_{F \cup Y}$ is an injective mapping into $F'' \cup Y''$, and $h|_F$ is a field homomorphism of F into F''. The implication (b) of Sublemma 4 shows that $h|_{F \cup Y}$ is a homomorphism of the bipartite graph (F, E, Y) into (F'', E'', Y''). So the restriction $h|_{F \cup Y}$ is a pre-morphism. By

$$\left(h(t(b, y))\right)^q = h\left((t(b, y))^q\right) = h(y-b) = h(y)-h(b) = \left(t(h(b), h(y))\right)^q,$$

we have $\left(h(t(b, y))/t(h(b), h(y))\right)^q=1$ for $b \in B(y)$.

Both $h(t(b, y))$ and $t(h(b), h(y))$ belong to R'', which clearly implies that they are equal. Now we prove that $h({}_iy)={}_i(h(y))$ for $i \in \omega$. We proceed by induction on i. The case $i=0$ is clear.

$$\left(h({}_{i+1}y)\right)^p = h\left(({}_{i+1}y)^p\right) = h({}_iy) = {}_i(h(y)) = \left({}_{i+1}(h(y))\right)^p,$$

therefore $\left(h({}_{i+1}y)/{}_{i+1}(h(y))\right)^p=1$. The quotient of two different elements from R'' is never a p^{th}-root of unity, so $h({}_{i+1}y)={}_{i+1}(h(y))$. Summarizing, we have proven that $h|_{F \cup R}$ is a pre-homomorphism.

(b) The uniqueness of the required extension is clear, since the set R generates the field extension $F(E, Y, p, q)/F$; further the restriction to $F \cup R$ of any possible extension must be a pre-homomorphism, and this pre-homomorphism is uniquely determined by the given pre-morphism. So the only problem is the existence of the extension.

Let K be the subfield of $F''(E'', Y'', p, q)$ generated by the range of the pre-homomorphism generated by the given pre-morphism. By the first part of the main lemma there is an isomorphism T from $F(E, Y, p, q)$ onto K, which is an extension of the given pre-morphism. Further, there exists the natural embedding U from K into $F''(E'', Y'', p, q)$. The composition TU is just the special homomorphism we need.

(c) The restriction of the special homomorphisms to the subset $F \cup Y$, as an operation, is an identity preserving and composition preserving bijection between the monoid of the special homomorphisms and that of the pre-morphisms.

The investigation of n-partite graphs

In the following $n>1$ denotes an integer. An $(n+1)$-tuple $(V_1, V_2, \ldots, V_n, E)$ will be called an n-partite graph iff the sets $V_1, V_2, \ldots, V_n$ are disjoint and E is a subset of the union of the direct products $V_i \times V_j$ where $1 \leqq i < j \leqq n$. Let V denote the union of the underlying sets $V_1, V_2, \ldots, V_n$, and let V'' be the union of the sets $V_1'', V_2'', \ldots, V_n''$. A mapping $f\colon V \to V''$ will be called a homomorphism of $(V_1, V_2, \ldots, V_n, E)$ into the n-partite graph $(V_1'', V_2'', \ldots, V_n'', E'')$ iff f is injective, V_i is mapped into V_i'' for $i=1, 2, \ldots, n$, and $\langle u, v\rangle \in E$ implies $\langle f(u), f(v)\rangle \in E''$. Let $\operatorname{Inj} \mathrm{PG}(n)$ denote the category whose objects are the n-partite graphs and whose morphisms are the homomorphisms defined above.

By the second part of the main lemma the structure of the monoid of the special endomorphisms is essentionally determined by the structure of the bipartite graph of the skin. If we iterate the special extension for different primes n times, then the special endomorphisms of that iterated extension can be described by the $(n+1)$-partite graph generated by the bipartite graphs of the skins.

Let N be a subset of $\{\langle i, j\rangle\colon 1 \leqq i < j \leqq n\}$. An n-partite graph $(V_1, V_2, \ldots, V_n, E)$ is said to be a unary n-partite graph of type N iff $E \cap (V_i \times V_j) = \emptyset$ for $\langle i, j\rangle \notin N$ and $(E \cap (V_i \times V_j))^{\mathrm{op}} = \{\langle v, u\rangle\colon u \in V_i,\ v \in V_j,\ \langle u, v\rangle \in E\}$ is a mapping of V_j into V_i for $\langle i, j\rangle \in N$. Let $\operatorname{Inj} \mathrm{UPG}(n, N)$ denote the full subcategory of $\operatorname{Inj} \mathrm{PG}(n)$ generated by the unary n-partite graphs of type N.

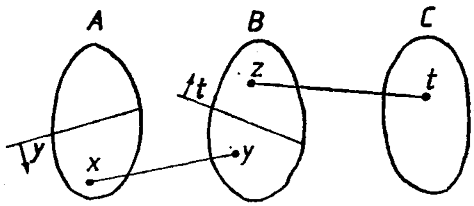
$\operatorname{Inj} \mathrm{PG}(n)$ cannot have a strong embedding into any category of algebras, as in the category $\operatorname{Inj} \mathrm{PG}(n)$ there are morphisms which are not isomorphisms while they are carried by injective mappings. In order to get a strongly algebraic subcategory of $\operatorname{Inj} \mathrm{PG}(n)$ which is "binding with respect to right-cancellative categories" we investigate the category $\operatorname{Inj} \mathrm{UPG}(n, N)$.

The following four claims offer a simple proof of the existence of arbitrary large rigid relational structures of bounded type (see P. Vopenka—A. Pultr—Z. Hedrlin [18]). The only fact which is not so trivial and will be used in the proof is that there are systems of almost disjoint sets which are large with respect to the underlying set. The n-partite graphs become simple relational structures by adding the unary relations $V_1, V_2, \ldots, V_n$. The n-partite graphs of the claims are rigid only with respect to their injective endomorphisms; but we can arrange that all the endomorphisms become injective by adding a further binary relation which is a full graph without loops.

Claim 1. *There exists a rigid 4-partite graph of cardinality* $(2^k)^+$, *if* k *is a strongly inaccessible cardinal.*

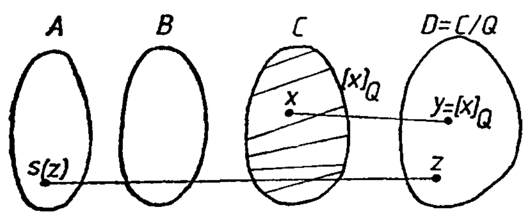
Proof. Let A be a set of cardinality k. As the cardinal k is strongly inaccessible there exists a set B of cardinality 2^k such that B is an almost disjoint system of subsets of A (see S. SHELAH [17]). This means that the cardinality of the intersection of any two different elements from B is less than the cardinality of A, while the cardinality of any element from B is equal to the cardinality of A. Further, there exists a set C of cardinality $(2^k)^+$ such that C is an almost disjoint system of subsets of B. Set

$$E'' = \{\langle x, y\rangle, \langle z, t\rangle: x\in A,\ y\in B,\ x\in y,\ z\in B,\ t\in C,\ z\in t\}.$$



Let the equivalence relation Q be the transitive hull of the following relation over C: two elements, u and v, are in relation iff there exist endomorphisms g and h of (A, B, C, E'') such that $g(u)=h(v)$. It is clear from the construction that each endomorphism of the 3-partite graph (A, B, C, E'') is determined by the action on the elements of A. So (A, B, C, E'') has at most 2^k endomorphisms. Therefore the cardinality of the factor set $D=C/Q$ is $(2^k)^+$. Let us fix a surjective mapping $s\colon D\to A$. Set

$$E = E''\cup\{\langle s(z), z\rangle, \langle x, y\rangle: z\in D, x\in C, y\in D, x\in y\}.$$

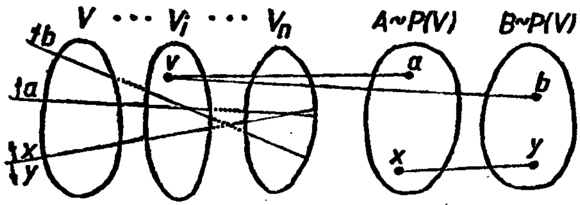


Since each of the equivalence classes, induced by the relation Q, is mapped into itself by any endomorphism of (A, B, C, E''), therefore each element of D is fixed by the endomorphisms of the 4-partite graph (A, B, C, D, E). So we get that (A, B, C, D, E) is a rigid 4-partite graph, and it is of cardinality $(2^k)^+$, as it was stated.

Claim 2. *There exists a rigid $(n+2)$-partite graph of cardinality 2^k, if there exists a rigid n-partite graph of cardinality $k\geqq\omega$.*

Proof. Let $(V_1, V_2, \ldots, V_n; E)$ be a rigid n-partite graph. Take two disjoint copies A and B of the power set of V. Set

$$E'' = E \cup \{\langle v, a\rangle, \langle v, b\rangle, \langle x, y\rangle: v \in V,\ a \in A,\ b \in B,\ v \in a,\ v \in b,$$
$$\text{and } V \text{ is the disjoint union of } x \text{ and } y \text{ where}$$
$$x \in A \text{ and } y \in B\}.$$

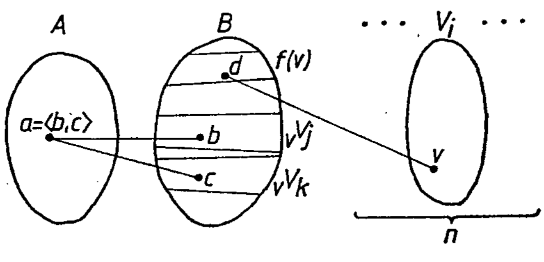


It is clear that $(V_1, V_2, \ldots, V_n; A, B, E'')$ is a rigid $(n+2)$-partite graph of cardinality $2^{|V|}$, if $|V| \geqq \omega$.

Claim 3. *Let* $(V_1, V_2, \ldots, V_n; E)$ *be a rigid n-partite graph of cardinality* $k \geqq \omega$. *Let further* $({}_vV_1; {}_vV_2, \ldots, {}_vV_{n_v}, {}_vE)$ *be a rigid* n_v*-partite graph of cardinality* k_v *for* $v \in V$. *Then there exists a rigid* $(n+2)$*-partite graph of cardinality* $\sum_{v \in V} k_v$.

Proof. Set $A = \bigcup\{{}_vE: v \in V\}$, $M = \{{}_vV_i: v \in V,\ 1 \leqq i \leqq n_v\}$ and $B = \bigcup M$. $|V| = |M|$ since $|V|$ is infinite. Let us fix a bijection $f: V \to M$. Set

$$E'' = E \cup \{\langle a, b\rangle, \langle a, c\rangle, \langle d, v\rangle: a \in A,\ b, c, d \in B,\ v \in V,\ a = \langle b, c\rangle,\ d \in f(v)\}.$$



Clearly $(A, B, V_1; V_2, \ldots, V_n, E'')$ is a rigid $(n+2)$-partite graph of cardinality $\sum_{v \in V} k_v$.

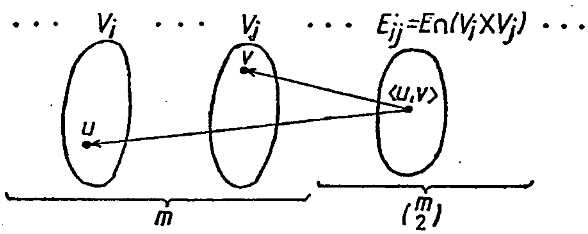
Claim 4. *For each cardinal number k there exists a natural number n such that there exists a rigid n-partite graph of cardinality greater than k.*

Proof. The proof is indirect. Let $k=\inf\{k'': \text{each rigid } n\text{-partite graph has less than } k'' \text{ elements for arbitrary } n\}$. By Claim 1 k is greater than $(2^{\omega})^{+}$, and k must be a regular strong limit by Claims 3 and 2. Further, k cannot be a strongly inaccessible cardinal by Claim 1.

Claim 5. *For each cardinal number k there exists a natural number n and a type N such that there exists a rigid unary n-partite graph of type N having cardinality greater than k.*

Proof. Let $(V_1; V_2, \ldots; V_m, E)$ be an arbitrary m-partite gaph. Set $E_{ij}= =E\cap(V_i\times V_j)$ for $1\leqq i<j\leqq m$. Further, let

$$E''=\{\langle u,\langle u,v\rangle\rangle,\ \langle v,\langle u,v\rangle\rangle:\ \langle u,v\rangle\in E\}.$$



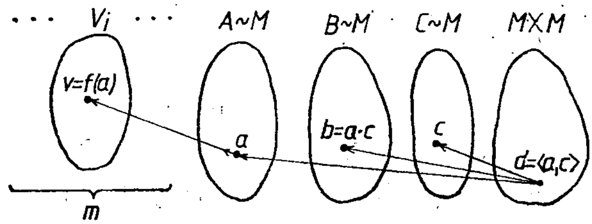
$(V_1; V_2, \ldots; V_m; E_{12}\ldots E_{1m}; E_{23}\ldots E_{2m}, \ldots, E_{(m-1)m}; E'')$ is a unary $\left(m+\binom{m}{2}\right)$-partite graph having the same endomorphism monoid as $(V_1, V_2, \ldots, V_m, E)$. Claim 4 finishes the proof.

Proposition 1. *Let M be a right cancellative monoid. Then there exists a natural number n such that there exists a unary n-partite graph such that its endomorphism monoid is isomorphic to M.*

Proof. By Claim 5 we can take a rigid unary m-partite graph $(V_1; V_2, \ldots, V_m; E)$ such that $|V|\geqq|M|$, and V is infinite. So we can fix an injective mapping $f\colon M\to V_i$ where the index i is suitably chosen. Now we take three disjoint copies A, B and C of the set M, and let

$$E''=E\cup\{\langle v,a\rangle,\ \langle a,d\rangle,\ \langle b,d\rangle,\ \langle c,d\rangle:\ v\in V,\ a\in A,\ b\in B,$$
$$c\in C,\ d\in M\times M,\ v=f(a),\ b=a\cdot c,\ d=\langle a,c\rangle\}.$$

An isomorphism between M and the endomorphism monoid of the unary $(m+4)$-partite graph $(V_1, V_2, \ldots, V_n, A, B, C, M\times M, E'')$ can be constructed on the basis of the following arguments.



Let h be any endomorphism of the $(m+4)$-partite graph. The construction yields that any element of the set $V\cup A$ is fixed by h. Suppose that $h(e)=c$, where e denotes the unit element of M being in the copy C. So the action of h on the set B is nothing else but the right multiplication by the element c. Further, the element $\langle x, y\rangle\in M\times M$ has to be mapped into the element $\langle x, y\cdot c\rangle$. Consequently h acts on the set C as the right multiplication by c.

Proposition 2. *For each similarity type t there is a natural number n and a type N such that there exists an extension of* $\operatorname{Inj Rel}(t)$ *into* $\operatorname{Inj UPG}(n, N)$.

Proof. Let $t\colon W\to$ Ordinal Numbers be a given similarity type. By Claim 5 there is a rigid unary m-partite graph $(V_1, V_2, \dots, V_m, E)$ such that $|V|\geqq|W|$, $|V|\geqq t_w$ for $w\in W$, and V is infinite. Clearly there is an index i such that we can fix an injective mapping $f\colon W\to V_i$ and injective mappings $f_w\colon t_w\to V_i$ for $w\in W$. Now we define an extension F of $\operatorname{Inj Rel}(t)$ into $\operatorname{Inj UPG}(m+3, N)$ for a suitable type N. Let (S, R) be a relational structure of similarity type t. This means that R_w is a t_w-ary relation over S for $w\in W$. Let

$$A=\bigcup\{\{w\}\times R_w\colon w\in W\},\quad B=\bigcup\{\{w\}\times R_w\times t_w\colon w\in W\}$$

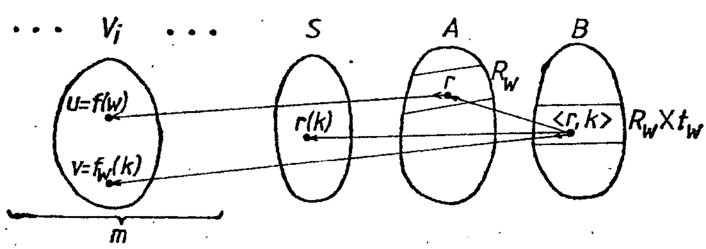
and

$$E''=E\cup\{\langle u, r\rangle, \langle r, b\rangle, \langle v, b\rangle, \langle s, b\rangle\colon u\in V,\ v\in V,\ s\in S,$$

$$r\in A,\ b\in B,\ u=f(w),\ v=f_w(k),\ b=\langle r, k\rangle, \text{ where } r\in R_w,\ k\in t_w$$

$$\text{and } s \text{ is the } k^{\text{th}} \text{ component of } r\}.$$

Let the unary $(m+3)$-partite graph $(V_1, V_2, \dots, V_m, S, A, B, E'')$ be the image of (S, R) under F.



Let h: $(S, R)\to(S'', R'')$ be an injective homomorphism. We define the morphism $F(h)$: $F(S, R)\to F(S'', R'')$ as follows:

$$F(h)_V = \mathrm{id}|_V, \quad F(h)|_S = h,$$
$$F(h)|_{\{w\}\times R_w} = \mathrm{id}|_{\{w\}}\times h^{(t_w)}|_{R_w} \quad \text{for} \quad w\in W,$$
$$F(h)|_{\{w\}\times R_w\times t_w} = \mathrm{id}|_{\{w\}}\times h^{(t_w)}|_{R_w}\times \mathrm{id}|_{t_w} \quad \text{for} \quad w\in W.$$

The operation F is clearly an extension if we can show that the functor F is full. In more detail it is enough to prove that for each morphism g: $F(S, R)\to F(S'', R'')$ the restriction $g|_S$ is a morphism of Inj Rel (t) and $F(g|_S)=g$. As $(V_1, V_2, ..., V_m; E)$ is rigid, $g|_V=\mathrm{id}|_V$. So the subset $\{w\}\times R_w$ of A is mapped into the subset $\{w\}\times R''_w$ of A''. Similarly the subset $\{w\}\times R_w\times t_w$ of B is mapped into the subset $\{w\}\times R''_w\times t_w$ of B''. Further we see that

$$g|_{\{w\}\times R_w\times t_w} = g|_{\{w\}\times R_w}\times \mathrm{id}|_{t_w}$$

and

$$g|_{\{w\}\times R_w} = \mathrm{id}|_{\{w\}}\times g^{(t_w)}|_{R_w} \quad \text{for} \quad w\in W.$$

Consequently, $g|_S$ is a homomorphism of (S, R) into (S'', R''). The remaining part of the proof is trivial.

Proposition 3. *For each similarity type t there is a natural number n and a type N such that there exists a strong embedding of* Inj Alg (t) *into* Inj UPG (n, N).

Proof. Let t: $W\to$Ordinal Numbers be a given similarity type. By Claim 5 there is a rigid unary m-partite graph $(V_1, V_2, ..., V_m, E)$ of type M and there is an index i such that there is an injective mapping s: $W\to V_i$ and there are injective mappings s_w: $t_w\to V_i$ for $w\in W$. Now we define a functor H: SET$\to$SET. For an arbitrary set A let

$$B_A = \bigcup\big\{\{w\}\times A^{(t_w)}\colon\ w\in W\big\} \quad \text{and} \quad C_A = \bigcup\big\{\{w\}\times A^{(t_w)}\times t_w\colon\ w\in W\big\}.$$

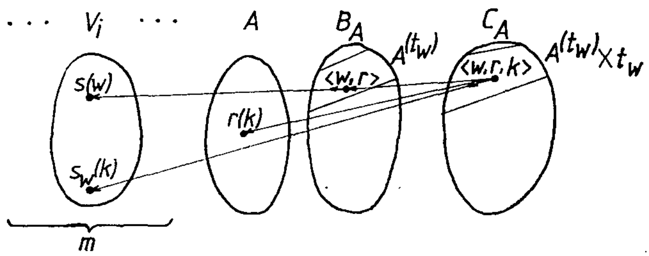
Let $H(A)=V\cup A\cup B_A\cup C_A$. If h: $A\to A''$ is a mapping between sets, then the mapping $H(h)$: $H(A)\to H(A'')$ is defined as follows:

$$H(h)|_V = \mathrm{id}|_V, \quad H(h)|_A = h,$$
$$H(h)|_{B_A} = \bigcup\,\{\mathrm{id}|_{\{w\}}\times h^{(t_w)}\colon\ w\in W\},$$
$$H(h)|_{C_A} = \bigcup\,\{\mathrm{id}|_{\{w\}}\times h^{(t_w)}\,\mathrm{id}|_{t_w}\colon\ w\in W\}.$$

Obviously, the functor H defined above is faithful. Set $n=m+3$ and $N=M\cup \cup\{\langle i, m+2\rangle, \langle m+1, m+2\rangle, \langle i, n\rangle, \langle m+1, n\rangle, \langle m+2, n\rangle\}$. Now we define a strong embedding F of Inj Alg (t) into Inj UPG (n, N) carried by the faithful functor H. Let (A, F) be an algebra of type t. This means that F_w is a t_w-ary operation over A for $w\in W$. Set

$$E'' = E\cup\{\langle s(w), \langle w, r\rangle\rangle, \langle F_w(r), \langle w, r\rangle\rangle, \langle s_w(k), \langle w, r, k\rangle\rangle,$$
$$\langle r(k), \langle w, r, k\rangle\rangle, \langle\langle w, r\rangle, \langle w, r, k\rangle\rangle \text{ where } \langle w, r. k\rangle\in C_A$$
$$\text{and } r(k) \text{ is the } k^{\text{th}} \text{ component of } r\}.$$

Let the unary n-partite graph $(V_1; V_2, ..., V_m, A, B_A, C_A; E'')$ of type N be the image of (A, F) under F. The underlying functor H uniquely determines the action of F on the morphisms.



To show that the functor F defined above is full the only non-trivial step is to prove the fact that F is a strong embedding, which is similar to the proof of Proposition 2.

The constructions

Main lemma (third part). *Let F be a given field of characteristic zero. Let $n>1$ be an integer and N be a type. Then there exists a strong embedding of* Inj UPG (n, N) *into* Ext $(F$, Fields). (The definitions can be found before Theorem 3 and before the Claim 1.)

Proof. By the Claim 5 there is a rigid unary k-partite graph $(W_1, W_2; ..., W_k; U)$ and an injective mapping s: $F_0\to W_j$ for a fixed index j where F_0 denotes the algebraic closure of F, and $|W_j|>|F_0|$. Let us further fix a sequence $r, q, p_0, p_1, p_2, ...$ $..., p_i, ...$ of different primes. Now we are able to define the underlying functor of the desired strong embedding. In order to avoid the complicated notations we define only the strong embedding, and later we give a simple argument to show that the strong embedding must be carried by a faithful endofunctor of SET.

The functor G: Inj UPG $(n, N)\to$Ext $(F$, Fields) is defined as follows. Let $(V_1, V_2, ...; V_n, E)$ be a unary n-partite graph. We define an ascending chain of fields $F_0\leqq F_1\leqq F_2\leqq...\leqq F_i\leqq...; i\in\omega$, by induction on i. Let $F_{i+1}=F_i(E_i, Y_i; p_i, q)$, where the bipartite graphs (F_i, E_i, Y_i) are the following:

$$\begin{cases} Y_0 = \{x\}, \\ E_0 = \emptyset; \end{cases}$$

$$\begin{cases} Y_i = W_i, \\ E_i = U\cap(F_i\times W_i), \end{cases} \quad \text{for} \quad i = 1, 2, \ldots, k;$$

$$\begin{cases} Y_{k+1} = F_0' \text{ where } F_0' \text{ is a new copy of } F_0, \\ E_{k+1} = \{\langle f\cdot x, f'\rangle, \langle s(f), f'\rangle \colon f\in F_0 \text{ and } f' \text{ corresponds to } f\}; \end{cases}$$

$$\begin{cases} Y_{k+1+i} = V_i, \\ E_{k+1+i} = E\cap(F_{k+1+i}\times V_i), \end{cases} \quad \text{for} \quad i = 1, 2, \ldots, n;$$

$$\begin{cases} Y_{k+n+2+2i} = W_{ji} \text{ where } W_{ji} \text{ is a new copy of } W_j, \\ E_{k+n+2+2i} = \{\langle w, w_i\rangle \colon w_i\in W_{ji} \text{ corresponds to } w\in W_j\}, \end{cases} \quad \text{for} \quad i\in\omega;$$

$$\begin{cases} Y_{k+n+3+2i} = W_j\times R_i \text{ where } R_i \text{ is the set of the roots of the } i^{\text{th}} \text{ skin} \\ E_{k+n+3+2i} = \{\langle r, \langle w, r\rangle\rangle, \langle w_i, \langle w, r\rangle\rangle \colon r\in R_i,\ w_i\in W_{ji} \text{ corresponds to } w\in W_j\} \end{cases}$$

$$\text{for } i\in\omega.$$

We define the image of $(V_1, V_2, \ldots, V_n, E)$ under G as the union of the above defined ascending chain of fields. Let $h\colon (V_1, V_2, \ldots, V_n, E)\to(V_1'', V_2'', \ldots, V_n'', E'')$ be a morphism from the category Inj UPG (n, N). We define an ascending chain of homomorphisms $h_0\leqq h_1\leqq\cdots\leqq h_i\leqq\ldots$, $i\in\omega$, where $h_i\colon F_i\to F_i''$. Let h_i be the identity of F_i for $i=0, 1, 2, \ldots, k+2$. Using the second part of the main lemma we get a unique extension h_{k+2+i} of h_{k+1+i} such that $h_{k+2+i}|_{V_i}=h|_{V_i}$ for $i=1, 2, \ldots, n$. Using the second part of the main lemma again we get an extension $h_{k+n+3+2i}$ of $h_{k+n+2+2i}$ and an extension $h_{k+n+4+2i}$ of $h_{k+n+3+2i}$ such that $h_{k+n+3+2i}$ is the identity on $Y_{k+n+2+2i}$ and $h_{k+n+4+2i}|_{Y_{k+n+3+2i}}=\mathrm{id}|_{Y_{k+n+2+2i}}\times h_{i+1}|_{R_i}$. Let, finally, $G(h)$ be the union of the ascending chain of the homomorphisms defined above. G is clearly a functor and an embedding. So we have to prove that G is full and a strong embedding.

By the first part of the main lemma, if an element w is either r-high or p_i-high in a field $G(V_1, V_2, \ldots, V_n, E)$, then either $w\in F_0$ or $w\in(F_{i+1}\setminus F_i)\cup F_0$, respectively. The subfield F_i can be defined as the set of those elements of $G(V_1, V_2, \ldots, V_n, E)$ which are algebraic over the set of s-high elements where s runs over $\{r, p_0, \ldots, p_{i-1}\}$. Therefore, each homomorphism of $G(V_1, V_2, \ldots, V_n, E)$ into $G(V_1'', V_2'', \ldots, V_n'', E'')$ maps the subfield F_i into the subfield F_i'', for $i\in\omega$. Consider the subset $\{\langle w, r\rangle \colon w\in W_j\}$ of $Y_{k+n+3+2i}$ for arbitrary given $r\in R_i$. The cardinality of this set is greater than F_0 and each element of this set is $p_{k+n+3+2i}$-high. Therefore, at least one of these variables cannot be mapped into the subfield $F_{k+n+3+2i}''$. Using the implication (b) of Sublemma 4 for this variable, we get that the subset R_i is mapped into the subset R_i'', for $i\in\omega$. Summarizing, we have proven that the restriction of each homomorphism of $G(V_1, V_2, \dots, V_n, E)$ into $G(V_1'', V_2'', \dots, V_n'', E'')$ to the subfield F_{i+1} is a special homomorphism of $F_i(E_i, Y_i, p_i, q)$ into $F_i''(E_i'', Y_i'', p_i, q)$. So we may use the second part of the main lemma for the subfields $F_i(E_i, Y_i, p_i, q)$, for $i\in\omega$. An obvious combinatorial argument finishes the proof of the fullness.

Now we prove that G is a strong embedding. Let $F(E, Y, p, q)$ and $F''(E'', Y'', p, q)$ be two special extensions such that the additive groups of F and F'' are isomorphic, and all the sets $A(y)$ and $A''(y'')$ have the same cardinality. The first part of the main lemma gives that each mapping of Y into Y'' naturally induces a group homomorphism of the additive group of $F(E, Y, p, q)$ into the additive group of $F''(E'', Y'', p, q)$. As the n-partite graphs contained in the constructed fields are unary n-partite graphs of a fixed type, the iteration of the above argument gives that each mapping of the underlying set of a unary n-partite graph into another one naturally induces a group homomorphism between the additive groups of the corresponding fields. This is, however, a much stronger property than that the embedding G is strong.

Proof of Theorem 2. Combining the third part of the main lemma and Proposition 1 we get the theorem.

Proof of Theorem 1. It follows from Theorem 2, as the one-element monoid is right cancellative.

Proof of Theorem 3. Using the third part of the main lemma and Propositions 2 and 3 we arrive at Theorem 3.

Proof of Theorem 4. The implications (c)⇒(b) and (b)⇒(a) are obvious, it is enough to prove that (a)⇒(c). By the fundamental theorem of binding categories (a review of the results can be found in the textbook of A. Pultr—V. Trnková [15]) it is enough to give a strong embedding of the category of 2-unary algebras into the category $\mathrm{Ext}\,(A, \mathrm{Alg}\,(t))$ whenever A is an algebra of similarity type t having no one-element subalgebra.

Let $A=(X, m, \dots)$ where m is an at least binary operation. In the following the polynomial $m(x, y, \dots, y)$ will be denoted simply by multiplication: $xy= =m(x, y, \dots, y)$. Now take a set Y disjoint to X such that $|Y|>|X|$, and $|Y|\geqq 8$. Let us fix an injective mapping $i\colon X\to Y$. Let (Y, R) be a rigid, connected, undirected graph having no loops (for the existence of such a graph see P. Vopenka—A. Pultr—Z. Hedrlin [18]). Take two further copies X_1 and X_2 of X, where $x_1\in X_1$ and $x_2\in X_2$ denotes the element corresponding to $x\in X$. Let us take three further elements: u, v and w not belonging to $Z\cup X\cup X_1\cup X_2\cup Y$.

Now we define a faithful endofunctor H of the category SET. For an arbitrary set Z let $H(Z)$ be the disjoint union of the sets Z, X, X_1, X_2, Y, $\{u\}$, $\{v\}$, and $\{w\}$. For an arbitrary mapping $h\colon Z\to Z''$ let $H(h)$ be the extension of h to $H(Z)$ acting identically on $H(Z)\setminus Z$.

Finally, we define a strong embedding F of the category of the 2-unary algebras into the category $\operatorname{Ext}(A, \operatorname{Alg}(t))$ such that the carrier of F is H. Let $(Z; g, h)$ be an arbitrary 2-unary algebra. Let the underlying set of $F(Z; g, h)$ be $H(Z)$. Recall, that xy denotes $m(x, y, \ldots, y)$. The operations of $F(Z; g, h)$ are defined as follows: Let $A=(X, m, \ldots)$ be a subalgebra of $F(Z; g, h)$. For $y\in Y$ and $b, c\in Y$, $b\neq c$, set

$$yy = u,$$
$$bc = b \quad\text{and}\quad cb = c \quad\text{if}\quad \langle b, c\rangle\in R,$$
$$bc = c \quad\text{and}\quad cb = b \quad\text{if}\quad \langle b, c\rangle\notin R.$$

For $x\in X$ (and for the corresponding $x_1\in X_1$ and $x_2\in X_2$) set

$$x_1x_2 = x, \quad x_2x_1 = i(x),$$

where $i\colon X\to Y$ is defined before.

$$uu = v, \quad vv = w.$$

For $z\in Z$ set

$$zv = u, \quad uz = g(z), \quad vz = h(z).$$

Otherwise the polynomial $m(x, y, \ldots, y)$ is defined by

$wq = v$ if $q\in H(Z)$, and the value of wq has not been defined yet,
$pq = w$ if $p, q\in H(Z)$, $p\neq w$, and the value of pq has not been defined yet.

In all the remaining cases let m be the projection to the first variable. All the operations are the projections to the first variable on the places where they haven't been defined yet.

The action of F on the morphisms is uniquely determined by the underlying functor H, which completes the definition of the functor F.

F is clearly a functor, an embedding and carried by H; so the only non-trivial property to prove is that F is full. This can be proved in the following nine steps:

(1) There is no one-element subalgebra of $F(Z; g, h)$ by the conditions on A and by the definition of m.

(2) Each two-element subset of Y is a subalgebra, consequently the restriction of any homomorphism of $F(Z; g, h)$ to Y is always injective.

(3) $(\exists c(b=bc))\Rightarrow b\in X\cup Y$, therefore Y is always mapped into $X\cup Y$ by any homomorphism.

(4) There must be a $y\in Y$ such that y is mapped into Y, since $|X|<|Y|$; consequently u is fixed, for $u=yy$ for all $y\in Y$.

(5) v and w are fixed together with u.

(6) Y is mapped into itself since it can be defined as the collection of those elements whose square is equal to u with respect to the multiplication.

(7) Y is mapped into itself in such a way that it is an injective strong endomorphism of the rigid graph (Y, R), by the definition of the multiplication. Therefore the set Y is fixed elementwise by any homomorphism.

(8) For $x \in X$, $x_2 x_1 = i(x)$; therefore either the images of x_2 and x_1 belong to Y, and consequently the image of x also belongs to Y, or the elements x_2 and x_1 are fixed, and consequently the element x is also fixed. Thus, each $x \in X$, and therefore each product xx is in $X \cup Y$, consequently none of the elements of X can go into Y. Therefore the set X is mapped into itself. This means that only the second case is possible: the sets X, X_1, and X_2 are fixed elementwise.

(9) Z is mapped into itself, since Z can be defined as the collection of those elements s for which $sv = u$. This mapping of Z is also an endomorphism of the 2-unary algebra $(Z; g, h)$ because of the definition of the multiplication by u and by v.

Hence the proof of Theorem 4 is finished.

References

[1] M. E. Adams—J. Sichler, Homomorphisms of bounded lattices with a given sublattice, *Arch. Math. (Basel)*, **30** (1978), 122—128.

[2] L. Babai—J. Nešetřil, High chromatic rigid graphs. I, in: *Combinatorics* (Proc. Conf. Keszthely, 1976), Colloq. Math. Soc. J. Bolyai, vol. 18, North-Holland (Amsterdam, 1978); pp. 53—60.

[3] E. Fried, Automorphism group of integral domains fixing a given subring, *Algebra Universalis*, **7** (1977), 373—387.

[4] E. Fried, Some properties of the category of integral domains, *Acta Math. Hung.*, **41** (1983), 3—15.

[5] E. Fried—J. Kollár, Automorphism groups of fields, in: *Universal algebra* (Proc. Conf. Esztergom, 1977), Colloq. Math. Soc. János Bolyai, vol. 29, North-Holland (Amsterdam, 1982); pp. 293—303.

[6] E. Fried—J. Sichler, Homomorphisms of commutative rings with unit element, *Pacific J. Math.*, **45** (1973), 485—491.

[7] E. Fried—J. Sichler, Homomorphisms of integral domains of characteristic zero, *Trans. Amer. Math. Soc.*, **225** (1977), 163—182.

[8] G. Grätzer, *Universal Algebra*, 2nd edition, Springer-Verlag (Berlin—Heidelberg—New York, 1979).

[9] J. De Groot, Groups represented by homeomorphism groups. I, *Math. Ann.*, **138** (1959), 80—102.

[10] J. Kollár, Some subcategories of integral domains, *J. Algebra*, **54** (1978), 329—331.

[11] J. Kollár, The category of unary algebras, containing a given subalgebra. I, *Acta Math. Acad. Sci. Hungar.*, **33** (1979), 407—417.

[12] J. KOLLÁR, The category of unary algebras, containing a given subalgebra. II, *Acta Math. Acad. Sci. Hungar.*, **35** (1980), 53—57.

[13] W. KUYK, The construction of fields with infinite cyclic automorphism group, *Canad. J. Math.*, **17** (1965), 665—688.

[14] S. MACLANE, *Categories for the Working Mathematician*, Springer-Verlag (Berlin, 1971).

[15] A. PULTR—V. TRNKOVÁ, *Combinatorial Algebraic and Topological Representations of Groups, Semigroups and Categories*, Academia (Prague, 1980).

[16] L. RÉDEI, *Algebra*, Pergamon Press (New York, 1967).

[17] S. SHELAH, *Classification Theory*, North-Holland (Amsterdam, 1978).

[18] P. VOPENKA—A. PULTR—Z. HEDRLIN, A rigid relation exists on any set, *Comment. Math. Univ. Carolinae*, **6** (1965), 149—155.

[19] B. L. VAN DER WAERDEN, *Algebra*, Springer-Verlag (Berlin, 1960).

DEPARTMENT OF ALGEBRA AND NUMBER THEORY
L. EÖTVÖS UNIVERSITY
MÚZEUM KRT. 6—8
1088 BUDAPEST, HUNGARY

Acta Sci. Math., 50 (1986), 39—49

Polynomial characterization of some idempotent algebras

J. DUDEK

0. Introduction. Let $\mathfrak{A}=(A, F)$ be an algebra. By $p_n=p_n(\mathfrak{A})$ we shall denote the number of all essentially n-ary polynomials over $\mathfrak{A}$. We say that two algebras (A, F_1) and (A, F_2) are polynomially equivalent if $A(F_1)=A(F_2)$ where $A(F_i)$ $(i=1, 2)$ denotes the set of all polynomials of (A, F_i). For such algebras we shall write $(A, F_1)=(A, F_2)$ (for details see [20], [30] and [31]).

In many cases the sequence $p_n=p_n(\mathfrak{A})$ determines the structure of the algebra $\mathfrak{A}$. For example, if $p_0=p_0(\mathfrak{A})=0$ and $p_n=p_n(\mathfrak{A})=1$ for all $n\geqq 1$, then $\mathfrak{A}$ is a nontrivial semilattice. For further nontrivial examples see, e.g., [21], [22], [27] and [43]. The following papers are concerned with $\{p_n\}$-sequences: [3], [4], [5], [21], [22], [23], [24], [25], [27], [28], [41], [42], [43], and [44].

An algebra $\mathfrak{A}=(A, F)$ is called proper if all $f\in F$ are pairwise distinct and every nonnullary $f\in F$ depends on its all variables. It is clear that an idempotent and commutative algebra $(A, +, \cdot)$ of type (2,2) (commutative means that both $+$ and $\cdot$ are commutative) is proper if and only if the polynomials $+$ and $\cdot$ are distinct. Denote by $V(+, \cdot)$ the variety of all idempotent and commutative algebras $(A, +, \cdot)$ of type $(2, 2)$. Recall that an algebra $(B, +, \cdot)$ of type $(2, 2)$ is a bisemilattice if both reducts $(B, +)$ and $(B, \cdot)$ are semilattices (for details see [32], [33], [38]). Denote by $B(+, \cdot)$ the variety of all bisemilattices. Of course, $B(+, \cdot)$ is a subvariety of $V(+, \cdot)$.

Now we shall present some results concerning the above varieties.

1. A polynomial characterization of lattices. We have

Theorem 1.1 (cf. [8], Theorem 1). *Let* $(B, +, \cdot)$ *be a bisemilattice with* $\operatorname{card} B\geqq 2$. *Then* $(B, +, \cdot)$ *is a lattice if and only if* $p_2(B, +, \cdot)=2$.

We should mention here that the assumption that $(B, +, \cdot)$ is a bisemilattice is essential. In fact, in [16] examples of idempotent commutative and nonassociative

Received November 28, 1983.

groupoids $(G, +)$ are given such that the polynomial $xy=(x+y)+y$ is a semilattice polynomial and in addition $p_2(G, +)=2$. Now treating this groupoid as a proper algebra $(G, +, \cdot)$ from $V(+, \cdot)$ we get our requirement since $p_2(G, +, \cdot)=p_2(G, +)=2$ and $(G, +, \cdot)$ is not a lattice.

Theorem 1.2 (see [8]). *There are no bisemilattices* $(B, +, \cdot)$ *for which* $p_2(B, +, \cdot)=3$.

Let us add that bisemilattices with $p_2=4$ are described in [17]. However the problem of finding all natural numbers m for which there exists a bisemilattice $(B, +, \cdot)$ such that $p_2(B, +, \cdot)=m$ is open. Recently JAN GAŁUSZKA [19] gave an example of a bisemilattice with $p_2=6$ and he also described all bisemilattices with 5 essentially binary polynomials.

It is a well-known fact that any proper distributive lattice $(L, +, \cdot)$ (i.e., distributive lattice with card $L \geqq 2$) has 9 essentially ternary polynomials, namely: $x+y+z$, xyz, $(x+y)z$, $(y+z)x$, $(z+x)y$, $xy+z$, $yz+x$, $zx+y$ and $xy+yz+zx$. As the next theorem shows this situation is exceptional for proper algebras from $V(+, \cdot)$.

Theorem 1.3. [12] *Let* $(A, +, \cdot)$ *be a proper algebra from* $V(+, \cdot)$. *Then* $(A, +, \cdot)$ *is a distributive lattice if and only if* $p_3(A, +, \cdot)=9$.

We also have the following

Theorem 1.4. [15] *Let* $(B, +, \cdot)$ *be a bisemilattice. Then* $(B, +, \cdot)$ *is a non-distributive modular lattice if and only if* $p_3(B, +, \cdot)=19$.

The proof of this theorem is rather complicated. It seems to be probable that this theorem would be true for proper algebras from the class $V(+, \cdot)$. This conjecture has been stated as a problem during the Problem Session of the Klagenfurt Conference on Universal Algebra (June, 1982).

2. Idempotent commutative groupoids. Let $V(\cdot)$ be the variety of all idempotent commutative groupoids $(G, \cdot)$. Recall that a groupoid $(G, \cdot) \in V(\cdot)$ satisfying $(xy)y=x$ is called a Steiner quasigroup (see [1]). Also recall that a groupoid $(G, \cdot)$ is distributive if $(xy)z=(xz)(yz)$ and $z(xy)=(zx)(zy)$ hold for all $x, y, z \in G$ (i.e. the right- and the left-distributive laws hold in $(G, \cdot)$).

In [14] the following theorem has been proved.

Theorem 2.1. *Let* $(G, \cdot)$ *be a proper groupoid from* $V(\cdot)$. *Then* $(G, \cdot)$ *is a distributive Steiner quasigroup if and only if* $p_3(G, \cdot)=3$.

A groupoid $(G, \cdot)$ will be called a near-semilattice (an upper bound algebra in the terminology of R. PARK [34]) if $(G, \cdot)$ is idempotent commutative and

$(xy)y=xy$, i.e., a subgroupoid of $(G, \cdot)$ generated by any two elements is a semilattice.

The following can be found in [14].

Theorem 2.2. *Let* $(G, \cdot)$ *be an idempotent groupoid with* $\operatorname{card} G \geqq 2$. *Then* $p_2(G, \cdot)=1$ *if and only if* $(G, \cdot)$ *is a Steiner quasigroup or a near-semilattice.*

Let m be an odd positive integer. Take an abelian group $(G, +)$ of exponent m. Denote by $G(m)$ the groupoid $(G, ((m+1)/2)(x+y))$. Using the result of [40] we infer that $G(m)$ is polynomially equivalent to the full idempotent reduct of the group $(G, +)$, i.e., $G(m)=(G, I(G, +))$, where $I(\mathfrak{A})$ denotes the set of all idempotent polynomials over $\mathfrak{A}$. If p is prime, then the groupoid $G(p)$ is called an affine groupoid and it is clear that $G(p)$ is polynomially equivalent to an affine space over the Galois field $\mathrm{GF}(p)$. Of course, all groupoids $G(m)$ are idempotent commutative distributive (even medial, i.e., $(xy)(uv)=(xu)(yv)$ holds in them) and the groupoid $G(3)$ is a medial Steiner quasigroup. For such groupoids we have also (see [2]):

$$p_n(G(m)) = ((m-1)^n-(-1)^n)/m \quad \text{for all } n.$$

In the case $m=3$ we have

$$p_n(G(3)) = (2^n-(-1)^n)/3 \quad \text{for all } n.$$

G. Grätzer and R. Padmanabhan [22] have proved the following:

Theorem 2.3 ([22], Theorem 3). *Let* $\mathfrak{A}=(A, \circ)$ *be an idempotent groupoid satisfying* $p_n(\mathfrak{A})=(2^n-(-1)^n)/3$ *for* $n=2, 3$ *and* 4. *Then a binary operation* $+$ *can be defined on* A *such that*

(i) $(A, +)$ *is an abelian group of exponent* 3;

(ii) *for all* $a, b \in A$, *we have* $a \circ b=2a+2b$.

In [13] the following generalization of the above is given:

Theorem 2.4. *Let* $(G, \cdot)$ *be a proper idempotent groupoid. Then* $(G, \cdot)$ *is an affine space over* GF (3) *(up to polynomial equivalence) if and only if* $p_4(G, \cdot)=5$.

A theorem similar to the above one is the following:

Theorem 2.5. [8] *Let* $(G, f(x_1, x_2, x_3, x_4))$ *be a proper (idempotent) symmetric algebra satisfying* $f(x, y, y, y)=x$. *Then* (G, f) *is polynomially equivalent to an affine space over* GF (3) *if and only if* $p_4(G, f)=5$.

3. Idempotent (noncommutative) groupoids. In this section we present two theorems. The proof of the first one can be found in [7]. The proof of the second theorem will be given in the last section.

Before formulating these theorems we need some notations. By Σ_0 we denote the variety of all semilattices, i.e., of all idempotent, commutative semigroups. Further notations are taken from [20] (pp. 394—395). By Σ_i $(i=1, 2, 3)$ we denote the varieties of groupoids $(G, \cdot)$ defined by the following identities:

$$\Sigma_1\colon\ x^2 = x,\quad (xy)z = x(yz),\quad x(yz) = x(zy),$$

$$\Sigma_2\colon\ x^2 = x,\quad (xy)z = (xz)y,\quad x(yz) = xy,\quad (xy)y = xy,$$

$$\Sigma_3\colon\ x^2 = x,\quad (xy)z = (xz)y,\quad x(yz) = xy,\quad (xy)y = x.$$

We say that a groupoid $(G, \circ)$ is a dual groupoid to a given $(G, \cdot)$ if $x\circ y=yx$ for all $x, y\in G$. If K is a class of groupoids, then K' denotes the class of all dual groupoids from K. Further, we put $T_i=\mathrm{HSP}(\{0, 1\}, \cdot_i)$, where $x_1\cdot_i x_2=x_i$ $(i=1, 2)$; i.e., T_i denotes the variety of all left (right) zero semigroups. The variety of all rectangular bands (see e.g. [36]) will be denoted by Δ. Recall that an idempotent semigroup $(G, \cdot)$ (i.e., a band) is called rectangular if $(G, \cdot)$ satisfies $xyz=xz$ for all $x, y, z\in G$. We should mention here that all the above groupoids were considered by many authors; see, e.g., [6], [18], [26], [29], [35], [41] and [42].

It is clear that

(i) $\Sigma_0\subset\Sigma_1$, $T_1\subset\Sigma_i$ for $i=1, 2, 3$ and $T_i\subset\Delta$ for $i=1, 2$.

We also have

(ii) $\Sigma_0\subset\Sigma_1$, $T_1'=T_2\subset\Sigma_i'$ $(i=1, 2, 3)$ and $\Delta=\Delta'$.

Recall that in [41] (see also [42]) it is shown that if a groupoid $(G, \cdot)$ satisfies $p_n(G, \cdot)=n$ for all n, then $(G, \cdot)$ belongs to the variety Σ_i or Σ_i' for some $i=1, 2, 3$. It is also not difficult to prove that if a groupoid $(G, \cdot)$ is neither a semilattice nor a singular semigroup, and $(G, \cdot)$ belongs to one of the classes Σ_i or Σ_i' $(i=1, 2, 3)$, then $p_n(G, \cdot)=n$ for all n. Thus we shall call a groupoid $(G, \cdot)$ an n-polynomial groupoid if $p_n(G, \cdot)=n$ for all n. Now Theorem 2 of [7] can be reformulated as follows:

Theorem 3.1. *Let $(G, \cdot)$ be an idempotent distributive groupoid having at most two essentially binary polynomials. Then $(G, \cdot)$ is either a semilattice or a Steiner quasigroup or a rectangular band or an n-groupoid or a noncommutative Steiner quasigroup (or its dual).*

Now we present the last theorem of this paper with complete proof.

Theorem 3.2. *Let $(G, \cdot)$ be an idempotent groupoid. Then $p_3(G, \cdot)<6$ if and only if $(G, \cdot)$ is either a semilattice or a distributive Steiner quasigroup or a rectangular band or an n-polynomial groupoid.*

4. Lemmas. In this section we formulate and prove all lemmas needed to prove the above theorem.

Lemma 4.1. *If $(G, \cdot)$ is a proper idempotent groupoid for which $x(yz)$ is not essentially ternary, then $(G, \cdot)$ satisfies either $x(yz)=xy$ or $x(yz)=xz$ (the dual version of the lemma is also true).*

Proof. Since $(G, \cdot)$ is proper we infer that the fundamental polynomial xy is essentially binary. Using this fact and the identity $x(yy)=xy$ we infer that $x(yz)$ depends on x. Now the assertion follows from the idempotency of xy.

Lemma 4.2. *If $(G, \cdot)$ is an idempotent groupoid which is not a rectangular band and satisfies the identity $x(yz)=xz$, then $p_3(G, \cdot)\geqq 6$.*

Proof. Since $(G, \cdot)$ is not a rectangular band we infer that $(G, \cdot)$ is proper and $(xy)z\neq xz$. If $(xy)z$ is not essentially ternary, then using the previous lemma (the dual version) we get $(xy)z=yz$. Hence using $x(yz)=xz$ we get $xy=(xy)(xy)= =y(xy)=yy=y$, a contradiction. Thus the polynomial $(xy)z$ is essentially ternary and therefore also all the following six polynomials are essentially ternary:

$$(*) \qquad (xy)z, \quad (yz)x, \quad (zx)y, \quad (yx)z, \quad (zy)x \quad \text{and} \quad (xz)y.$$

Now we prove that all the polynomials of $(*)$ are pairwise distinct. We give here only the proof for the inequality $(xy)z\neq(yx)z$ (for the other inequalities the proof runs similarly). Assume $(xy)z=(yx)z$. Putting uv instead of x in this identity we get $((uv)y)z=(y(uv))z=(yv)z$. Thus we infer that both polynomials $((x_1x_2)x_3)x_4$ and $x_1(x_2(x_3x_4))$ are not essentially 4-ary. This contradicts Lemma 3 of [3]. Hence $p_3(G, \cdot)\geqq 6$.

Lemma 4.3. *If $(G, \cdot)$ is a proper idempotent groupoid satisfying $x(yz)=xy$, then the polynomial $(xy)z$ is essentially ternary and the inequalities $(xy)z\neq(yz)x$, $(xy)z\neq(zx)y$, $(xy)z\neq(yx)z$ and $(xy)z\neq(zy)x$ hold in $(G, \cdot)$.*

Proof. If $(xy)z$ is not essentially ternary, then using Lemma 2 of [3] we infer that $(G, \cdot)$ is a rectangular band. Hence $xy=x(yz)=xz$ which gives $xy=x$, a contradiction. We prove only that $(xy)z=(yx)z$ does not hold in $(G, \cdot)$. (Analogously one can prove the remaining inequalities.) Assume that $(xy)z=(yx)z$. Then $ux=u(xy)=u((xy)z)=u((yx)z)=u(yx)=uy$. This proves that $(G, \cdot)$ is improper, a contradiction.

Lemma 4.4. *If $(G, \cdot)$ satisfies $x^2=x$, $x(yz)=xy$ and $(xy)z=(xz)y$, then $(G, \circ)$, where $x\circ y=(xy)y$ satisfies the same identities, i.e., $x\circ x=x$, $x\circ(y\circ z)= =x\circ y$ and $(x\circ y)\circ z=(x\circ z)\circ y$.*

Proof. We check only the last identity. We have $(x\circ y)\circ z=(((xy)y)z)z=$ $=(((xz)z)y)y=(x\circ z)\circ y$.

Lemma 4.5. *If* $(G, \cdot)$ *satisfies the assumption of the previous lemma and* $x\circ y=(xy)y$ *is not essentially binary, then* $(G, \cdot)\in\Sigma_3$.

Proof. Since $(G, \cdot)$ is idempotent we infer that $(xy)y=x$ or $(xy)y=y$. If the first case holds, then $(G, \cdot)\in\Sigma_3$. If the second identity holds, then we get $xy=$ $=x((xy)y)=x(xy)=xx=x$. This gives $y=(xy)y=xy=x$ and hence $(G, \cdot)\in\Sigma_3$ as a one-element algebra.

Lemma 4.6. *If* $(G, \cdot)$ *satisfies the assumption of Lemma* 4.4 *and* $\operatorname{card} G\geqq 2$, *then* $(xy)y\neq yx$.

Proof. Assume that $(xy)y=yx$. Then we get $xy=x(yx)=x((xy)y)=x(xy)=$ $=xx=x$. This gives $x=xy=(xy)y=yx=y$, a contradiction.

Lemma 4.7. *If* $(G, \cdot)$ *satisfies the assumption of Lemma* 4.4, $x\circ y=(xy)y$ *is essentially binary, then* $(G, \cdot)$ *belongs to* Σ_2 *or* $p_3(G, \cdot)\geqq 6$.

Proof. If $x\circ y=xy$, then $(G, \cdot)\in\Sigma_2$. Further, using Lemmas 4.5 and 4.6 we can assume that $x\circ y$ is essentially binary and $x\circ y\notin\{xy, yx\}$. First of all we shall check that $(G, \circ)$ is noncommutative. Applying $x(yz)=xy$ we get the requirement. Indeed, if $x\circ y=y\circ x$, then $x=xx=x(xy)=x((xy)y)=x(x\circ y)=$ $=x(y\circ x)=x((yx)x)=x(yx)=xy$, a contradiction. Consider now the algebra $(G, \cdot, \circ)$ of type $(2, 2)$ and its ternary polynomials

$$(**)\qquad (xy)z,\quad (yz)x,\quad (zx)y,\quad (x\circ y)\circ z,\quad (y\circ z)\circ x\quad\text{and}\quad (z\circ x)\circ y.$$

Using Lemmas 4.1, 4.4, and the fact that both $\cdot$ and $\circ$ are essentially binary we verify that these polynomials are essentially ternary. Further observe that the polynomial $(xy)z$ does not admit any permutation of its variables except the identity permutation and the transposition (y, z) (the same is true for the polynomial $(x\circ y)\circ z$). In fact, if the converse is true, then $(xy)z$ is a symmetric polynomial and hence $xy=(xx)y=(yx)x=((yy)x)x=((yx)y)x=(yx)(yx)=yx$. Hence $xy=$ $=x(yz)=x(zy)=xz=x$, a contradiction. Now using the fact that $(xy)z$ and $(x\circ y)\circ z$ admit only the transposition (y, z) and the identity permutation of its variables we infer that all polynomials of $(**)$ are pairwise distinct. Let us add that to prove this fact we also use the noncommutativity of xy, $x\circ y$ and $x\circ y\notin\{xy, yx\}$. Thus $p_3(G, \cdot)\geqq 6$.

Lemma 4.8. *If* $(G, \cdot)$ *is an idempotent groupoid satisfying* $x(yz)=xy$ *and* $p_3(G, \cdot)<6$, *then* $(G, \cdot)$ *is either a left zero semigroup or* $(G, \cdot)$ *is an n-polynomial groupoid.*

Proof. If $(G, \cdot)$ is improper, then $xy=x$ or $xy=y$. The first identity proves that $(G, \cdot)\in T_1$ and the second one gives that $y=yy=y(xy)=yx=x$ which proves that $(G, \cdot)$ also is a left zero semigroup. Assume now that $(G, \cdot)$ is proper. Using Lemma 4.3 we infer that the polynomial $(xy)z$ is essentially ternary. If $(xy)z\neq\neq(xz)y$, then using Lemma 4.3 we get $p_3(G, \cdot)\geqq 6$, a contradiction. Assume now that $(G, \cdot)$ satisfies the identity $(xy)z=(xz)y$. Consider the binary polynomial $x\circ y=(xy)y$. If $x\circ y$ is not essentially binary, then applying Lemma 4.5 we infer that $(G, \cdot)$ is an n-polynomial groupoid since $(G, \cdot)\in\Sigma_3$. If $x\circ y=xy$, then $(G, \cdot)\in\Sigma_2$ and hence $(G, \cdot)$ is an n-polynomial groupoid. If $x\circ y$ is essentially binary and $x\circ y\neq xy$, then applying Lemma 4.7 we get $p_3(G, \cdot)\geqq 6$, a contradiction. The proof of the lemma is completed.

Lemma 4.9. *Let $(G, \cdot)$ be an idempotent groupoid. Then $(G, \cdot)$ is a semilattice if and only if $(xy)z=y(zx)$ holds in $(G, \cdot)$.*

Proof. The necessity is obvious. Assume that $(xy)z=y(zx)$ holds identically; then

$$yx=y(xx)=(xy)x=x(yx)=(xx)y=xy,\quad (xy)z=(yx)z=x(zy)=x(yz),$$

.e., $(G, \cdot)$ is commutative and associative, as required.

Lemma 4.10. *If $(G, \cdot)$ is an idempotent noncommutative nonassociative groupoid with both polynomials $(xy)z$ and $x(yz)$ essentially ternary, then the groupoid $(G, \cdot)$ contains at least six essentially ternary polynomials.*

Proof. First of all observe that $(G, \cdot)$ is proper. Indeed, if $(G, \cdot)$ is improper, then $xy=x$ or $xy=y$ and hence $(G, \cdot)$ is associative which contradicts the assumption. Using Theorem 8 of [8] we infer that the polynomials $(xy)z$, $(yz)x$ and $(zx)y$ are pairwise distinct. The same we have for the polynomials $x(yz)$, $y(zx)$ and $z(xy)$. Now consider the following six essentially ternary polynomials over $(G, \cdot)$:

$$(xy)z,\quad (yz)x,\quad (zx)y,\quad x(yz),\quad y(zx)\quad \text{and}\quad z(xy).$$

Using the previous lemma, the noncommutativity and the nonassociativity of the fundamental polynomial xy we infer that the above polynomials are pairwise distinct which finishes the proof of the lemma.

Lemma 4.11. *Let $(G, \cdot)$ be a band. Then $p_3(G, \cdot)<6$ if and only if $(G, \cdot)$ is either a rectangular band or $(G, \cdot)$ satisfies $xyz=xzy$ or $xyz=yxz$.*

Proof. Let $p_3(G, \cdot)<6$. Consider two cases (1) xyz is not essentially ternary and (2) xyz is essentially ternary. Take into consideration the first case. If xy is not essentially binary, then $(G, \cdot)$ is a singular semigroup and hence a rectangular band. Further assume that (1) holds and $(G, \cdot)$ is proper, then using Lemma 4.1

we infer that $xyz=xz$. Thus $(G, \cdot)$ is a rectangular band (the same follows from Lemma 2 of [3]). Assume now (2). Since $p_3(G, \cdot)<6$ we infer at least two among the essentially ternary polynomials xyz, yzx, zxy, yxz, zyx, xzy are equal. If xy is commutative, then xyz is a symmetrical polynomial and the assertion of the lemma is obvious. Now let $(G, \cdot)$ be noncommutative. In this case it is routine to prove that the considered groupoid satisfies $xyz=xzy$ or $xyz=yxz$. To prove the converse it suffices to observe that $p_3(G, \cdot)=0$ for any rectangular band (see [37]). Using results of [41] we infer that $p_n(G, \cdot)=n$ for all n and all $(G, \cdot)\in\Sigma_1\cup\Sigma_1'$ if $(G, \cdot)$ is neither a semilattice nor a singular semigroup. Thus for any groupoid $(G, \cdot)$ being a rectangular band or belonging to $\Sigma_1\cup\Sigma_1'$ we have $p_3(G, \cdot)\leqq 3$ since for a semilattice we have $p_n=1$ for all n (if $\operatorname{card} G\geqq 2$). This proves the lemma.

To prove the next lemma we need a result from [11]:

Lemma 4.12 (see Theorem 8 of [11]). *Let $(G, \cdot)$ be an idempotent commutative groupoid. Then $(G, \cdot)$ is a semilattice if and only if the polynomial $f(x, y, z)=$ $=(xz)(yz)$ is symmetric.*

Now we prove

Lemma 4.13. *Let $(G, \cdot)$ be an idempotent commutative groupoid. Then $p_3(G, \cdot)<6$ if and only if $(G, \cdot)$ is either a semilattice or $(G, \cdot)$ is a distributive Steiner quasigroup.*

Proof. Assume that $p_3(G, \cdot)<6$. If $(G, \cdot)$ is associative, then $(G, \cdot)$ is a semilattice. If $(G, \cdot)$ is nonassociative, then the groupoid $(G, \cdot)$ is proper and the polynomials $(xy)z$, $(yz)x$ and $(zx)y$ are essentially ternary and pairwise different. Take now the polynomial $f(x, y, z)=(xz)(yz)$. Since $(G, \cdot)$ is idempotent and commutative we infer that f is essentially ternary. Further observe that if f admits a permutation of its variables, except the transposition (x, y) and the identity permutation, then f is symmetric. Now applying Lemma 4.12 we infer that the polynomials $(xz)(yz)$, $(xy)(zy)$, $(yx)(zx)$ are different and, of course, essentially ternary. Consider now the following essentially ternary polynomials: $(xy)z$, $(yz)x$, $(zx)y$, $(xz)(yz)$, $(xy)(zy)$ and $(yx)(zx)$. Since $p_3(G, \cdot)<6$ we infer that at least two of these polynomials are equal. Using the commutativity and the nonassociativity of xy we deduce that $(G, \cdot)$ satisfies $(xy)z=(xz)(yz)$, i.e., $(G, \cdot)$ is distributive. Consider the polynomial $x\circ y=(xy)y$. Using Theorem 1 of [10] (see also [22]) we infer that $(xy)y\neq y$. If $x\circ y$ is not essentially binary, then $x\circ y=x$ and hence $(G, \cdot)$ is a distributive Steiner quasigroup. Assume now that $x\circ y$ is essentially binary. If $x\circ y=xy$ then using Theorem 8 of [10] we infer that $(G, \cdot)$ is a semilattice, a contradiction. If $x\circ y\neq xy$, then the proof splits into two cases: (1) $x\circ y$ is commutative, (2) $x\circ y$ is noncommutative. If the first case holds, then using again Theorem 8 of [10] we infer that $(G, \cdot)$ is a semilattice, a contradiction (see also

[39]). If (2) holds, then using the main result of [4] (see also [28]) we get $p_3(G, \cdot) \geqq 7$, a contradiction. This finishes the proof of the first part of the lemma. To prove the converse it suffices to observe that for any proper semilattice $(G, \cdot)$ we have $p_n(G, \cdot)=1$ for all n and $p_3(Q, \cdot)=3$ for any proper distributive Steiner quasigroup $(Q, \cdot)$ (see Theorem 2.1). Thus in both cases we have $p_3<6$. This completes the proof of the lemma.

5. **Proof of Theorem 3.2.** First of all we have $p_n=1$ $(n=1, 2, \ldots)$ for all proper semilattices, $p_3=3$ for a proper distributive Steiner quasigroup and $p_3 \leqq 3$ for any n-polynomial groupoid (see [41]). Hence for all these groupoids we have $p_3<6$. Assume now that $(G, \cdot)$ is idempotent and $p_3(G, \cdot)<6$. Now, if $(G, \cdot)$ is commutative, then the proof follows from Lemma 4.13. If $(G, \cdot)$ is noncommutative, then we consider the polynomials $g_1(x, y, z)=(xy)z$ and $g_2(x, y, z)=x(yz)$. If none of g_1 and g_2 is essentially ternary, then using Lemma 2 of [3] we infer that $(G, \cdot)$ is a rectangular band. Further assume that g_1 is essentially ternary and g_2 is not essentially ternary. Then the proof follows from Lemmas 4.1, 4.2 and 4.8 (if g_2 is essentially ternary and g_1 is not essentially ternary then we apply the dual version of the previous lemmas). If both polynomials g_1 and g_2 are essentially ternary, then the proof follows from Lemmas 4.10 and 4.11. Thus the proof of Theorem 3.2 is completed.

References

[1] W. D. Belousov, *Introduction to the theory of quasigroups and loops* (Moscow, 1967). (in Russian)

[2] B. Csákány, On affine spaces over prime fields, *Acta Sci. Math.*, **37** (1975), 33—36.

[3] J. Dudek, Number of algebraic operations in idempotent groupoids, *Colloq. Math.*, **21** (1970), 169—177.

[4] —, Binary minimal algebras, Lectures on the theory of ordered sets and general algebra (Abstracts, Summer Session, Harmonia, 1970), *Acta Fac. Rerum Natur. Univ. Comenian. Math. Mimoriadne Čislo*, **1971**, 21—22.

[5] —, On non-associative groupoids, *Colloq. Math.*, **36** (1976), 23—25.

[6] —, A new characterization of groupoids with at most n essentially n-ary polynomials, *Bull. Soc. Roy. Sci. Liège*, **48** (1980), 390—392.

[7] —, Some remarks on distributive groupoids, *Czechoslovak Math. J.*, **31** (1981), 58—64.

[8] —, On bisemilattices. I, *Colloq. Math.*, **47** (1982), 1—5.

[9] —, An affine space over GF (3) as a 4-groupoid, in: *Proceedings of the Skopje Conference on n-ary Structures* (1982); pp. 243—251.

[10] —, On binary polynomials in idempotent commutative groupoids, *Fund. Math.*, **120** (1984), 187—191.

[11] —, Varieties of idempotent commutative groupoids, *Fund. Math.*, **120** (1984), 193—204.

[12] —, A characterization of distributive lattices, in: *Contributions to Lattice Theory* (Proc. Conf. Szeged, 1980), Colloquia Mathematica Societatis János Bolyai, 33, North-Holland (Amsterdam, 1983); pp. 325—335.

[13] —, A polynomial characterization of affine spaces over GF (3), *Colloq. Math.*, to appear.
[14] —, On Steiner quasigroups, *Math. Slovaca.*, to appear.
[15] —, A polynomial characterization of nondistributive modular lattice, *Colloq. Math.*, to appear.
[16] J. Dudek and K. Głazek, On commutative groupoids with semilattice operations, *Simon Stevin*, **55** (1981), 247—264.
[17] J. Dudek and A. Romanowska, Bisemilattices with four essentially binary polynomials, in: *Contributions to Lattice Theory*, (Proc. Conf. Szeged, 1980), Colloquia Mathematica Societatis János Bolyai, 33, North-Holland (Amsterdam, 1983); pp. 337—359.
[18] T. Evans, Products of points — some simple algebras and their identities, *Amer. Math. Monthly*, **74** (1967), 362—372.
[19] J. Gałuszka, Bisemilattices with five essentially binary polynomials, preprint.
[20] G. Grätzer, *Universal Algebra*, 2nd ed., Springer-Verlag (Berlin, 1979).
[21] —, Compositions of functions, in: *Proc. Conf. on Universal Algebra*, Queen's University (Kingston, Ont., 1970); pp. 1—106.
[22] G. Grätzer and R. Padmanabhan, On idempotent commutative and nonassociative groupoids, *Proc. Amer. Math. Soc.*, **28** (1971), 75—80.
[23] G. Grätzer and J. Płonka, On the number of polynomials of an idempotent algebra. I, *Pacific J. Math.*, **32** (1970), 697—709.
[24] —, On the number of polynomials of an idempotent algebra. II, *Pacific J. Math.*, **47** (1973), 99—113.
[25] G. Grätzer, J. Płonka and A. Sekanina, On the number of polynomials of a universal algebra. I, *Colloq. Math.*, **22** (1970), 9—11.
[26] N. Kimura and M. Yamada, Note on idempotent semigroups. II, *Proc. Japan. Acad.*, **34** (1958), 110—112.
[27] A. Kisielewicz, The p_n-sequences of idempotent algebras are strictly increasing, *Algebra Universalis*, **13** (1981), 233—250.
[28] K. M. Koh, Idempotent algebras with three essentially binary polynomials, *Algebra Universalis*, **10** (1980), 232—246.
[29] E. S. Ljapin, *Semigroups* (Moscow, 1960). (in Russian)
[30] E. Marczewski, Independence and homomorphisms in abstract algebras, *Fund. Math.*, **50** (1961), 45—61.
[31] —, Independence in abstract algebras. Results and problems, *Colloq. Math.*, **14** (1966), 169—188.
[32] R. McKenzie and A. Romanowska, Varieties of $\cdot$-distributive bisemilattices, in: *Contributions to general algebra* (Proc. Conf. Klagenfurt, 1978), Verlag Johannes Heyn (Klagenfurt, 1967); pp. 213—217.
[33] R. Padmanabhan, Regular identities in lattices, *Trans. Amer. Math. Soc.*, **158** (1971), 179—188.
[34] R. E. Park, A four element algebra whose identities are not finitely based, *Algebra Universalis*, **11** (1980), 255—260.
[35] F. Pastijn, Rectangular bands of inverse semigroups, to appear.
[36] M. Petrich, *Lectures in semigroups*, Akademie-Verlag (Berlin, 1977).
[37] J. Płonka, Diagonal algebras, *Fund. Math.*, **58** (1966), 309—321.
[38] —, On distributive quasilattices, *Fund. Math.*, **60** (1967), 197—200.
[39] —, On the number of independent elements in finite abstract algebras with two binary symmetrical operations, *Colloq. Math.*, **19** (1968), 9—21.
[40] —, On the arity of idempotent reducts of groups, *Colloq. Math.*, **21** (1970), 35—37.

[41] —, On algebras with n distinct n-ary operations, *Algebra Universalis*, **1** (1971), 73—79.
[42] —, On algebras with at most n distinct operations, *Algebra Universalis*, **1** (1971), 80—85.
[43] —, On the minimal extension of the sequence (0, 0, 2, 4), *Algebra Universalis*, **3** (1973), 335—340.
[44] K. URBANIK, On algebraic operations in idempotent algebras, *Colloq. Math.*, **13** (1965), 129—157.

UNIWERSYTET WROCŁAWSKI
INSTYTUT MATEMATYCZNY
PL. GRUNWALDZKI 2/4
50 384 WROCŁAW, POLAND

Bäcklund's theorem and transformation for surfaces V_2 in E_n

Ü. LUMISTE

1. Introduction and classical background. Bäcklund's classical transformation gives a way to generate new solutions of the Sine—Gordon equation $\partial^2\psi/\partial u^1\partial u^2 = \sin\psi$ from a given solution. Its geometrical setting uses some basic propositions for surfaces V_2 in E_3, which are the following.

A. If there is a diffeomorphism $V_2 \to V_2^*$, $x \mapsto x^*$, between two distinct surfaces in E_3 such that $\overrightarrow{xx^*} \in T_x V_2 \cap T_{x^*} V_2$, $|\overrightarrow{xx^*}| = r = \text{const}$ and the angle φ between $T_x V_2$ and $T_{x^*} V_2^*$ is a constant, then both V_2 and V_2^* have constant negative Gaussian curvature equal to $-(\sin^2\varphi)/r^2$ (the classical Bäcklund's theorem [1], [2]).

B. This diffeomorphism, called a pseudospherical line congruence, maps asymptotic curves of V_2 to asymptotic curves of V_2^* (i.e. it is a Weingarten congruence or W-congruence).

C. Asymptotic curves of a surface V_2 with Gaussian curvature $K = \text{const} < 0$ (i.e. of an immersion of a piece of the Bolyai—Lobachevsky plane $L_2(K)$ into E_3) form a Chebyshev net: in suitable net parameters u^1 and u^2 the metric of V_2 can be given by $ds^2 = (du^1)^2 + 2\cos\psi \cdot du^1\, du^2 + (du^2)^2$.

D. The net angle ψ of a Chebyshev net of a Riemannian V_2 satisfies the equation $\partial^2\psi/\partial u^1\partial u^2 = -K\sin\psi$, where K is the Gaussian curvature of V_2; in case if V_2 is a piece of $L_2(-1)$ this equation is the Sine—Gordon equation.

Due to C and D, every immersion of a piece of $L_2(-1)$ into E_3 gives a solution ψ of the Sine—Gordon equation and this correspondence is one to one up to rigid motion. Due to A and B, there is a transformation of such a solution to another, the analytical formulation of which gives Bäcklund's classical transformation [2].

The aim of this paper is to give some generalizations of propositions A, B and C to the case of surfaces V_2 in E_n, $n > 3$. Note that D needs no generalization because it does not depend on the immersion.

Received January 2, 1984 and in revised form August 17, 1984.

A generalization of the geometrical Bäcklund theorem and transformation in another direction, for the case of V_m in E_{2m-1}, is given in [4], [5]. If $m=2$ this reduces to the classical one.

2. Main results. The next generalization of Bäcklund's theorem gives some additions to the classical case too.

Theorem 1 [3]. *For two distinct surfaces V_2 and V_2^* in E_n, $n \geqq 3$, let $V_2 \to V_2^*$, $x \mapsto x^*$ be a diffeomorphism such that $\overrightarrow{xx^*} \in T_x V_2 \cap T_{x^*} V_2^*$ and $|\overrightarrow{xx^*}| = r \neq 0$ for every point $x \in V_2$. Let φ be the angle between $T_x V_2$ and $T_{x^*} V_2^*$, and let K and K^* be the Gaussian curvatures of V_2 and V_2^* in the corresponding points x and x^*. Then the following four conditions are equivalent:*

(1) $r = const$ and $\varphi = const$,
(2) $K = K^* = -(\sin^2 \varphi)/r^2 = const$,
(3) $K = K^* = -(\sin^2 \varphi)/r^2$ and $r = const$,
(4) $K = K^* = -(\sin^2 \varphi)/r^2$ and $\varphi = const$.

Under the assumptions of this theorem the diffeomorphism $V_2 \to V_2^*$ is called the line pseudocongruence (if $n=3$ "pseudo" is to be dropped); V_2 and V_2^* are called its focal surfaces. They cannot be arbitrary surfaces, but necessarily must consist of planar points only. Tangent planes $T_x V_2$ and $T_{x^*} V_2^*$ in corresponding points x and x^* lie in an Euclidean 3-plane $(E_3)_x$. Among the second fundamental tensors of V_2 in normal directions to $T_x V_2$ we can distinguish the tensor h in the normal direction lying in $(E_3)_x$. A pair of null directions of the tensor h is called a pair of h-asymptotic directions in $T_x V_2$ and corresponding curves on V_2 are called h-asymptotic curves. The diffeomorphism $V_2 \to V_2^*$ is called h-asymptotic if it maps h-asymptotic curves of V_2 to h-asymptotic curves of V_2^*.

Theorem 2. *Under the same assumptions as in Theorem 1 the next three conditions are equivalent to each other and also to each of the conditions* (1)—(4):

(5) $K = -(\sin^2 \varphi)/r^2 = const$ *and* $V_2 \to V_2^*$ *is h-asymptotic,*
(6) $K = -(\sin^2 \varphi)/r^2$, $r = const$ *and* $V_2 \to V_2^*$ *is h-asymptotic,*
(7) $K = -(\sin^2 \varphi)/r^2$, $\varphi = const$ *and* $V_2 \to V_2^*$ *is h-asymptotic.*

Here the Gaussian curvature K of V_2 can be replaced of course by the Gaussian curvature K^* of V_2^*.

Theorem 3. *Under the same assumptions as in Theorem 1 let one of the conditions* (1)—(7) *be satisfied (and hence each of them). Let the field of distinguished normal directions (i.e. belonging in each $x \in V_2$ to $(E_3)_x$) be parallel along the curves tangent to directions of $\overrightarrow{xx^*}$ with respect to normal connection of V_2. Then the net of h-asymptotic curves on V_2 is a Chebyshev net.*

Theorem 3, due to proposition D, gives a possibility to find a solution of the Sine—Gordon equation by the special immersion of a piece of the Bolyai—Lobachevsky plane $L_2(-1)$ into E_n. Theorems 1 and 2 show how this solution can be then transformed.

The well-known Hilbert's theorem [6] states, that there is no solution $\psi: R^2 \to R$ of the Sine—Gordon equation, which is different from 0 and π in every point $(u^1, u^2) \in R^2$. It follows, that the class of surfaces V_2, satisfying the assumptions of Theorem 3, does not include the Bolyai—Lobachevsky plane $L_2(-1)$, globally immersed into E_n with regular h-asymptotic net. That gives a contribution to the theorems about classes of surfaces V_2 in E_n, which does not contain a V_2 isometric with $L_2(-1)$ (see [7]).

Here it is important that a surface V_2, satisfying the assumptions of Theorem 3, can be defined by following conditions, without turning to $V_2^* \subset E_n$: 1) V_2 consists of planar points only, 2) the field of normal curvature directions, corresponding to the lines of conjugated net family of V_2, is parallel along the lines of the same family with respect to normal connection, 3) invariants r and φ (which can be expressed in terms of V_2 only) are constants and $r^2 = \sin^2 \varphi$. In this paper we cannot give the complete explanation of the question about impossibility to realize $L_2(-1)$ by such a V_2. It needs a new publication.

3. Frame restriction. A local field of orthonormal frames will be choosen so that the origin is $x \in V_2$ and $e_1, e_2 \in T_x V_2$. In formulae $dx = e_I \theta^I$, $de_I = e_K \theta_I^K$; $I, K, \ldots = 1, \ldots, n$; $d\theta^I = \theta^K \wedge \theta_K^I$, $d\theta_K^I = \theta_K^L \wedge \theta_L^I$, $\theta_I^K + \theta_K^I = 0$ for the field of orthonormal frames in E_n we have then $\theta^3 = \ldots = \theta^n = 0$, and hence $\theta^1 \wedge \theta_1^\alpha + \theta^2 \wedge \theta_2^\alpha = 0$; $\alpha, \beta, \ldots = 3, \ldots, n$. By Cartan's lemma we may write $\theta_i^\alpha = b_{ij}^\alpha \theta^j$, $b_{ij}^\alpha = b_{ji}^\alpha$, $i, j, \ldots = 1, 2$. From the assumptions of Theorem 1 it follows that the tangent planes $T_x V_2$ and $T_{x^*} V_2^*$ lie in an Euclidean 3-plane $(E_3)_x$ because $T_x V_2 \cap T_{x^*} V_2^* \ni \overrightarrow{xx^*} \neq 0$. The frame can be choosen so that $e_3 \in (E_3)_x$ in each point $x \in V_2$ and $e_1 = (1/r)\overrightarrow{xx^*}$. Then the point $x^* \in V_2^*$, corresponding to $x \in V_2$, has the radius vector $x^* = x + re_1$ and from

$$dx^* = (\theta^1 + dr)e_1 + (\theta^2 + r\theta_1^2)e_2 + r(\theta_1^3 e_3 + \theta_1^\varrho e_\varrho), \tag{3.1}$$

$$\varrho, \sigma, \ldots = 4, \ldots, n,$$

it follows that by such a choice of the frame we have $\theta_1^\varrho = 0$. Thus $b_{11}^\varrho = b_{12}^\varrho = 0$.

The linear span of normal curvature vectors $b_{ij}^\alpha X^i X^j e_\alpha$ with arbitrary unit vector $X^i e_i \in T_x V_2$ is called the first normal space $N_x^1 V_2$. Now it has dimension two because it is spanned on $b_{11} = b_{11}^3 e_3$, $b_{12} = b_{12}^3 e_3$ and $b_{22} = b_{22}^\alpha e_\alpha$, the first two of which are collinear. We can finally restrict our choice of the frame by the condition that $e_4 \in N_x^1 V_2$ in each point $x \in V_2$. Then

$$b_{22}^5 = \ldots = b_{22}^n = 0;$$

and so we have

$$\theta_i^3 = h_{ij}\theta^j, \quad h_{21} = h_{12}, \quad \theta_1^4 = 0, \quad \theta_2^4 = k_{22}\theta^2, \quad \theta_i^5 = \ldots = \theta_i^n = 0,$$

where the notations $h_{ij} = b_{ij}^3$ and $k_{22} = b_{22}^4$ are used.

4. Gaussian curvatures. The above restriction can be done for the surface V_2^* choosing the frame vectors e_I^* at the point $x^* \in V_2^*$ in a similar way. Then

$$\begin{aligned} e_1^* &= e_1, \\ e_2^* &= e_2 \cos\varphi + e_3 \sin\varphi, \\ e_3^* &= -e_2 \sin\varphi + e_3 \cos\varphi, \\ e_4^* &= e_4, \ldots, e_n^* = e_n, \end{aligned}$$

and

$$dx^* = e_1^*\theta^{*1} + e_2^*\theta^{*2} = e_1\theta^{*1} + (e_2\cos\varphi + e_3\sin\varphi)\theta^{*2}.$$

Comparing with (3.1) we have

$$\theta^{*1} = \theta^1 + dr, \quad \theta^{*2}\cos\varphi = \theta^2 + r\theta_1^2, \quad \theta^{*2}\sin\varphi = r\theta_1^3.$$

Here $\sin\varphi$ cannot be 0 because this would lead to $\theta_1^3 = 0$ and V_2 would be a torse with line generators xx^* and we had $V_2^* = V_2$ what is excluded by the assumptions of Theorem 1. Therefore

$$(1/r)\theta^2 + \theta_1^2 = \cot\varphi \cdot \theta_1^3.$$

From this, by exterior differentiation and using well-known formulae,

$$d\theta_1^2 = -K\theta^1 \wedge \theta^2, \quad \theta_1^3 \wedge \theta_2^3 = K\theta^1 \wedge \theta^2 \tag{4.1}$$

we have (see [3])

$$K = -((\sin^2\varphi)/r^2)(1 + r_1) + (h_{12}\varphi_1 - h_{11}\varphi_2), \tag{4.2}$$

where $dr = r_1\theta^1 + r_2\theta^2$, $d\varphi = \varphi_1\theta^1 + \varphi_2\theta^2$.

For the surface V_2^*,

$$\theta_1^{*3} = de_1^* \cdot e_3^* = ((\sin\varphi)/r)\theta^2, \quad \theta_2^{*3} = de_2^* \cdot e_3^* = \theta_2^3 + d\varphi,$$

and now the second formula in (4.1) for V_2^* gives ([3])

$$K^* = -\frac{\sin^2\varphi}{r^2}\,\frac{h_{12} + \varphi_1}{h_{12}(1 + r_1) - h_{11}r_2}. \tag{4.3}$$

These formulae (4.2) and (4.3) for the Gaussian curvatures K and K^* will be used in the proof of Theorems 1 and 2, but they also have their own significance.

5. Proof of Theorem 1. If $r=\text{const}$ and $\varphi=\text{const}$, then from (4.2) and (4.3) we obtain (2), (3) and (4) immediately. Conversely, let

$$K=K^*=-(\sin^2\varphi)/r^2.$$

Then the same formulae (4.2) and (4.3) give correspondingly

$$(5.1)\qquad h_{12}\varphi_1-h_{11}\varphi_2=((\sin^2\varphi)/r^2)r_1,$$
$$h_{12}r_1-h_{11}r_2=\varphi_1.$$

In case of (2) we have $K=\text{const}$ and from $Kr^2+\sin^2\varphi=0$ it follows that $dr=r\cot\varphi\cdot d\varphi$ and the last two equations give $\sin^2\varphi\cdot\varphi_1=0$. Therefore $\varphi_1=r_1=0$ and $h_{12}\varphi_2=h_{11}r_2=0$. Here $h_{11}=0$ would lead to $\varphi=0$ what is excluded, and we have (1).

In case of (2) or (3), when $r=\text{const}$ or $\varphi=\text{const}$, the same equations give (1). Theorem 1 is proved.

6. Proof of Theorem 2. The h-asymptotic curves of V_2 are the null curves of the second fundamental form in the direction e_3. This form is

$$\mathrm{II}^3=\theta^1\theta_1^3+\theta^2\theta_2^3=h_{ij}\theta^i\theta^j.$$

For V_2^* it is

$$\mathrm{II}^{*3}=\theta^{*1}\theta_1^{*3}+\theta^{*2}\theta_2^{*3}=(\theta^1+dr)\frac{\sin\varphi}{r}\theta^2+\frac{r}{\sin\varphi}\theta_1^3(\theta_2^3+d\varphi).$$

Using here that $h_{11}h_{22}-h_{12}^2=K$ and (4.2), we have

$$\mathrm{II}^{*3}=\frac{rh_{12}}{\sin\varphi}\mathrm{II}^3+\Phi,$$

where

$$\Phi=\frac{r\varphi_1}{\sin\varphi}[h_{11}(\theta^1)^2+2h_{12}\theta^1\theta^2]+\left(\frac{\sin\varphi}{r}r_2+\frac{r}{\sin\varphi}h_{12}\varphi_2\right)(\theta^2)^2.$$

If $r=\text{const}$ and $\varphi=\text{const}$, then $\Phi=0$, and we have (5). Conversely, let $V_2\to V_2^*$ be h-asymptotic. Then Φ must be proportional to II^3, and therefore

$$(6.1)\qquad h_{22}\varphi_1=\frac{\sin^2\varphi}{r^2}r_2+h_{12}\varphi_2.$$

In case of (5) we have $r_i=r\cot\varphi\cdot\varphi_i$. Now (5.1) and (6.1) give

$$\left(h_{12}-\frac{\sin 2\varphi}{2r}\right)\varphi_1-h_{11}\varphi_2=0,$$

$$h_{22}\varphi_1-\left(h_{12}+\frac{\sin 2\varphi}{2r}\right)\varphi_2=0.$$

Here the determinant is :

$$K+\frac{\sin^2 2\varphi}{4r^2} = -\frac{\sin^4\varphi}{r^2} \neq 0$$

and hence (1) holds.

In case of (6), from (5.1) and (6.1) it follows that

$$h_{12}\varphi_1 - h_{11}\varphi_2 = 0,$$
$$h_{22}\varphi_1 - h_{12}\varphi_2 = 0,$$

where the determinant is $-h_{12}^2+h_{11}h_{22}=K\neq 0$. In case of (7) it follows analogously that (1) holds. Theorem 2 is proved.

7. Proof of Theorem 3. If the field of directions e_3 is parallel along the integral curves of the equation $\theta^2=0$ with respect to normal connection, then

$$\theta_3^4 = \lambda\theta^2. \tag{7.1}$$

Taking the unit vectors

$$\hat{e}_1 = e_1\cos\alpha + e_2\sin\alpha, \quad \hat{e}_2 = -e_1\sin\alpha + e_2\cos\alpha$$

in h-principal directions, bisecting h-asymptotic directions we have $\hat{h}_{12}=0$ and besides this

$$\begin{aligned} \hat{\theta}_1^4 &= \theta_2^4 \sin\alpha = k_{22}\sin\alpha\cdot\theta^2, \\ \hat{\theta}_2^4 &= \theta_2^4\cos\alpha = k_{22}\cos\alpha\cdot\theta^2, \\ \hat{\theta}_i^\varrho &= 0. \end{aligned} \tag{7.2}$$

The local parameters v^1 and v^2 on V_2 can be choosen so that

$$\hat{\theta}^1 = a_1 dv^1, \quad \hat{\theta}^2 = a_2 dv^2, \quad \hat{\theta}_1^3 = b_1 a_1 dv^1, \quad \hat{\theta}_2^3 = b_2 a_2 dv^2,$$

where $b_1=\hat{h}_{11}$, $b_2=\hat{h}_{22}$. Then from the formulae

$$d\hat{\theta}_1^3 = \hat{\theta}_1^2\wedge\hat{\theta}_2^3, \quad d\hat{\theta}_2^3 = -\hat{\theta}_1^2\wedge\hat{\theta}_1^3,$$

which hold due to (7.1) and (7.2), using the well-known expression

$$\hat{\theta}_1^2 = \frac{1}{a_1}\frac{\partial a_2}{\partial v^1}dv^2 - \frac{1}{a_2}\frac{\partial a_1}{\partial v^2}dv^1,$$

we have

$$\frac{1}{b_i-b_j}\frac{\partial b_i}{\partial v^j} = -\frac{\partial(\ln a_i)}{\partial v^j}, \quad i\neq j.$$

The same computation as in [2] leads us to parameters w^1, w^2, in which for V_2

$$ds^2 = \cos^2\chi(dw^1)^2 + \sin^2\chi(dw^2)^2,$$
$$\mathrm{II}^3 = \sin\chi\cos\chi[(dw^1)^2-(dw^2)^2]$$

and now by $u^1=w^1+w^2$, $u^2=w^1-w^2$ we get

$$ds^2 = (du^1)^2+2\cos\psi\, du^1\, du^2+(du^2)^2,$$

$$\mathrm{II}^3 = 2\sin\psi\, du^1\, du^2.$$

The h-asymptotic net is a Chebyshev net. Theorem 3 is proved.

References

[1] A. V. Bäcklund, Om utor med konstant negativ krökning, *Lunds Univ. Arsskrift,* **19** (1883).

[2] S.-S. Chern, C.-L. Terng, An analogue of Bäcklund's theorem in affine geometry. *Rocky Mountain J. Math.,* **10** (1) (1980), 105—124.

[3] Ü. Lumiste, L. Tuulmets, On Gaussian curvatures of local surfaces V_2 of a line pseudocongruence in E_n, *Tartu Ülikooli Toimetised, Acta et Comm. Univ. Tartuensis,* **665** (1984), 55—62. (Russian)

[4] K. Tenenblat, C.-L. Terng, Bäcklund's theorem for n-dimensional submanifolds of R^{2n-1}, *Ann. Math.* **111** (1980), 477—490.

[5] C.-L. Terng, A higher dimensional generalization of the Sine—Gordon equations and its soliton theory, *Ann. Math.,* **111** (1980), 491—510.

[6] D. Hilbert, Über Flächen von konstanter Gausschen Krümmung, *Trans. Amer. Math. Soc.,* **2** (1901), 87—99.

[7] S. B. Kadomcev, Impossibility of some special isometric immersions of the Lobachevsky space, *Mat. Sbornik,* **107** (1978), 175—198. (Russian)

FACULTY OF MATHEMATICS
TARTU STATE UNIVERSITY
202400 TARTU, USSR

Acta Sci. Math., 50 (1986), 59—66

On cell complexes generated by geodesics in the non-Euclidean elliptic plane

SUREN V. APIKIAN

Abstract. In this paper we consider some properties of cell complexes in the non-Euclidean elliptic space, which are generated by n geodesic lines. The cells are geodesically convex polygons and in non-degenerate case the number of cells equals $\binom{n}{2}+1$. Each cell-complex has at least one cell with maximal number of vertices. If we denote by ξ_n this maximal number and $\alpha_n=\min \xi_n$, where the minimum is taken over all possible complexes, then we show that $\alpha_3=3$, $\alpha_4=4$ and $\alpha_n=5$ for all $n\geqq 5$.

Introduction. Consider the space G of straight lines on the plane R^2. Let $O\in R^2$ be the origin and denote by $[O]$ the bundle of lines through O. For $g\in G\setminus[O]$ let (p, φ) be the polar coordinates of the foot of perpendicular from the origin on g. It is usual to consider the pair (p, φ) as coordinates of the line g, where $p\in]0, \infty[$, $\varphi\in[0, 2\pi]$. Thus $G\setminus[O]$ is mapped onto semi-cylinder without rim, having ordinary cylindric coordinates. Note that diametrically opposite points on the rim correspond to the same line from the bundle $[O]$. Hence for the space G we obtain the model C of a semi-cylinder with identified opposite points on the rim (see A. Baddeley in [2]). By means of central projection the manifold C can be mapped onto the elliptic plane E_2 with punctured pole N (see Fig. 1).

We shall denote the corresponding homeomorphism by $\Phi\colon G\to E_2\setminus N$. Especially important is that under Φ the bundles of lines on R^2 correspond to the geodesics in E_2 (see [1]).

Note that the inverse mapping Φ^{-1} is well-defined if an origin and a reference direction are chosen.

Denote by Γ the space of geodesic lines γ in E_2. One can easily see that $\Phi^{-1}(\gamma)$ is either a bundle of parallels (if $N\in\gamma$), or a bundle of lines, passing through a point $\mathcal{P}\in R^2$. Thus we have the mapping $\Psi\colon \Gamma\setminus[N]\to R^2$. A collection of geodesics

Received January 6, 1984.

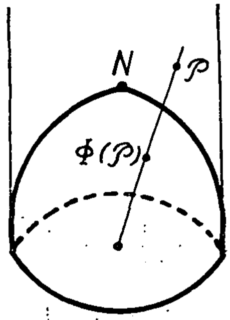


Figure 1. The points $\mathscr{P}$ and $\Phi(\mathscr{P})$ lie on a line through the centre of the circle.

$\{\gamma_i\}_{i=1}^n$ is called non-degenerate, if no three γ_i-s pass through a point. If we choose the pole $N\overline{\in}\bigcup_{i=1}^n \gamma_i$, then $\{\mathscr{P}_i=\Psi(\gamma_i)\}_{i=1}^n$ is a set of points in R^2. Note, that if $\{\gamma_i\}_{i=1}^n$ is non-degenerate, then the corresponding set $\{\mathscr{P}_i\}_{i=1}^n$ consists of n points in general position, i.e. no three lie on a line.

The application of mappings Φ^{-1} and Ψ facilitates the analysis of many properties of cell-complexes by reducing the problem to the investigation of the corresponding sets in G. The efficiency of such approach was suggested to the author by R. V. Ambartzumian.

The main problem. We consider a cell-complex in E_2, generated by non-degenerate collections of geodesic lines $\{\gamma_i\}_{i=1}^n$. Our problem is as follows. In non-degenerate case the number of cells equals $\binom{n}{2}+1$ and every cell-complex has at least one cell with maximal number of neighbours (the cells are polygons, they are regarded as neighbours, if they have a common side). If we denote this maximal number by ξ_n then the problem is to find $\min \xi_n$, where the minimum is with respect to all possible non-degenerate collections $\{\gamma_i\}_{i=1}^n$. Let us fix a non-degenerate $\{\gamma_i\}_{i=1}^n$ and consider the corresponding set $\{\mathscr{P}_i\}_{i=1}^n$. We call two lines $g_1, g_2 \in G$ $(g_1, g_2 \overline{\in} \bigcup_{i=1}^n [\mathscr{P}_i])$ equivalent if they produce the same separation of the set $\{\mathscr{P}_i\}_{i=1}^n$ into two subsets.

Further we shall call each class of equivalent lines an atom in G. Each atom corresponds (via Φ^{-1}) to a cell in E_2 and there is exactly one unbounded atom, which corresponds to the cell in E_2 containing the pole N. We shall use the following algorithm to determine the number of neighbours of a cell α from a cell-complex on E_2.

Algorithm. Denote by g_{ij} the straight line through the points $\mathscr{P}_i$ and $\mathscr{P}_j$. The number of neighbours of the atom α is equal to the number of lines from the collection g_{ij}, which belong to the boundary of the atom $\Phi^{-1}(\alpha)$ (the latter lines will be termed "limiting" lines of the atom).

For example (see Fig. 2) the atom containing the line g is a pentagon.

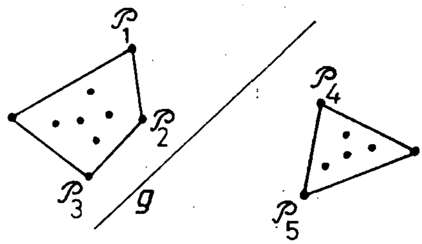


Figure 2. The limiting lines of the atom generated by g are $g_{12}, g_{23}, g_{15}, g_{34}, g_{45}$.

We state (without proof) the following simple lemma.

Lemma. *The number of sides of the minimal convex hull of the set* $\{\mathcal{P}_i\}_{i=1}^n$ *is equal to the number of neighbours of the unbounded atom.*

Let us make some remarks on the properties of cell-complexes. Here and below the term "cell-complex" (c.c.) will mean "partition of the non-Euclidean elliptic plane E_2 by a non-degenerate family of geodesics".

Remark I. Let $n \geqq 5$, and suppose that among the atoms of c.c. there is at least one n-gon. Then the c.c. consists of exactly one n-gon, n triangles and $n(n-3)/2$ quadrangles.

Proof. Let us denote the n-gon by α. We choose the pole N in α and construct the mapping Ψ: $\Gamma \diagdown [N] \to R^2$. By the Lemma $\{\mathcal{P}_i\}_{i=1}^n$ forms a convex n-gon.

Applying the above algorithm one can easily show (Fig. 3), that the remaining atoms are either triangles or quadrangles. Namely, the atoms, which contain a line

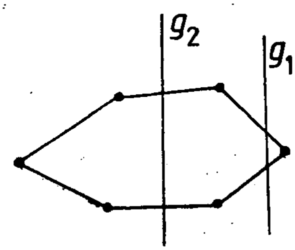


Figure 3. The line g_1 belongs to a triangular atom and g_2 belongs to a quadrangular atom.

separating one vertex from the others, are triangles and their number is exactly n. All other atoms are quadrangles and their number is $\binom{n}{2}-n=n(n-3)/2$.

Remark II. If $n>3$, then $\min \xi_n>3$.

Proof. Suppose, to the contrary, that there exists a c.c. consisting of triangles only. Then, by the Lemma, for any choice of N, the minimal convex hull of the set $\{\mathscr{P}_i=\Psi(\gamma_i)\}_{i=1}^n$ is a triangle.

It is not difficult to see that the atom defined by the line shown in Fig. 4 is a k-gon, with $k\geqq 4$. This contradiction proves Remark II.

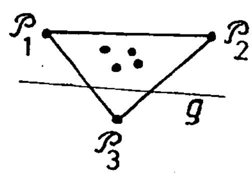


Figure 4. $\mathscr{P}_3$ is separated from all other points.

Remark III. If $n>4$, then $\min \xi_n>4$.

Proof. Suppose, to the contrary, that there exists a c.c. consisting of triangles and quadrangles only. By (II) we can find a quadrangular atom. If we choose the pole N in this atom then, by the Lemma, let the points $\mathscr{P}_1$, $\mathscr{P}_2$, $\mathscr{P}_3$ and $\mathscr{P}_4$ form the minimal convex hull of $\{\mathscr{P}_i\}_{i=1}^n$ (Fig. 5).

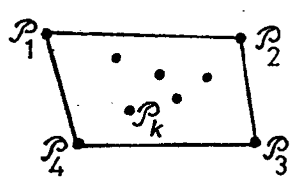


Figure 5.

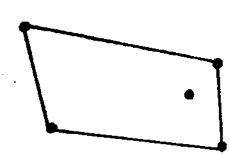


Figure 6.

Consider the collection $\{\mathscr{P}_i\}_{i=1}^n\diagdown\{\mathscr{P}_k\}$ (we delete the point $\mathscr{P}_k$), where $\mathscr{P}_k$ belongs to the interior of $\mathscr{P}_1\mathscr{P}_2\mathscr{P}_3\mathscr{P}_4$. This corresponds to the deletion of the geodesic γ_k on E_2. It can be proved that the deletion of this geodesic cannot result in the formation a new polygon with more than 4 sides. Deleting successively the points different from $\mathscr{P}_1$, $\mathscr{P}_2$, $\mathscr{P}_3$, $\mathscr{P}_4$, we obtain a five-point set (see Fig. 6). Here we can easily show the pentagonal atom. This contradiction proves Remark III.

Remark IV. Denote by q_k the number of k-gons of the c.c. The method described above answers the following question: What are the possible sequences $(q_1, q_2, \dots, q_n)$, generated by c.c.? We have found that for $n=3, 4, 5$ all possible cases are as follows:

$$n = 3, \quad q_3 = 4,$$
$$n = 4, \quad q_3 = 4, \quad q_4 = 5,$$
$$n = 5, \quad q_3 = 5, \quad q_4 = 5, \quad q_5 = 1.$$

For $n=6$ we have the following possibilities.

$$q_3 = 6, \quad q_4 = 9, \quad q_5 = 0, \quad q_6 = 1,$$
$$q_3 = 10, \quad q_4 = 0, \quad q_5 = 6, \quad q_6 = 0,$$
$$q_3 = 6, \quad q_4 = 8, \quad q_5 = 2, \quad q_6 = 0.$$

In particular, we obtain that $\min \xi_3=3$, $\min \xi_4=4$, $\min \xi_5=5$. What is the $\min \xi_n$, when $n>5$? The answer is given by the following

Theorem.

$$\min \xi_n = \begin{cases} n, & n \leqq 5 \\ 5, & n > 5. \end{cases}$$

Proof. It is sufficient to construct such a set of points on R^2, which have only "triangular", "quadrangular" or "pentagonal" atoms.

Consider a unit square $\mathcal{P}_1\mathcal{P}_2\mathcal{P}_3\mathcal{P}_4$ on R^2 (see Fig. 7).

We shall place the points $\{\mathcal{P}_i\}_{i=1}^n$ on congruent arcs σ_i, which emanate from vertices $\mathcal{P}_i$ $(i=1, 2, 3, 4)$ and lie within the square. Two points $Q_1 \in \sigma_i$ and $Q_2 \in \sigma_j$ $(i \neq j)$ are called corresponding, if Q_1 goes in Q_2 under euclidean motion, which brings σ_i in to σ_j.

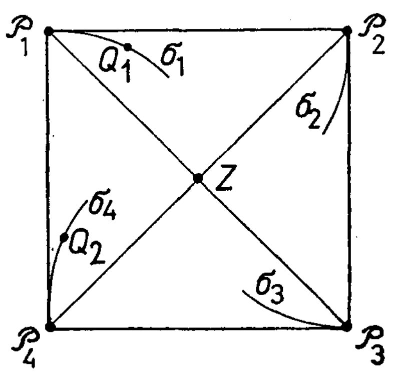


Figure 7.

Let $y=f(x)$ be an equation of σ_1. We shall find $f(x)$, using the following condition:

The tangent at every point of σ_1 crosses σ_2 in the point, which corresponds to the point of tangency on σ_1.

From this we derive the differential equation

$$\text{(A)} \qquad \frac{1-x-y}{y-x}=\frac{dy}{dx} \quad \text{with initial condition} \quad y(0)=1.$$

The solution of this Cauchy problem exists, it is unique and it is a logarithmic helix. From (A) we deduce that the curve $y=f(x)$ is convex in the neighbourhood of $\mathscr{P}_1$ and it has horizontal tangent at the point $\mathscr{P}_1$ (side of the square).

We take each σ_i to be a "piece" of logarithmic helix. Denote by λ the common length of the arcs σ_i $(i=1,2,3,4)$. Let $\delta_1, \delta_2, \ldots, \delta_k, \ldots$ be a sequence of positive numbers such that $\sum \delta_i<\lambda$. Now we proceed to construct the desired set. First we construct an auxiliary sequence of points $\{Q_i\}$ on the curve σ_1. Let Q_1 be the endpoint of σ_1 and if the points $Q_1, Q_2, \ldots, Q_j$ have been constructed, then Q_{j+1} is constructed as follows. We draw the line $\mathscr{P}_1 Q_j$. Let Q_j' be the intersection point of this line with σ_2 (see Fig. 8). Starting from Q_j' we move along σ_2 in the direction of $\mathscr{P}_2$ at distance δ_j. In this way we obtain the point Q_j''. Now we draw a line through Q_j'', which is tangent to σ_1 and let Q_{j+1} be the point of tangency. It is clear that in this way an infinite sequence of points $\{Q_i\}_{i=1}^{\infty}$ can be constructed. Now we describe how we construct the collection $\{\mathscr{P}_i\}_{i=1}^{n}$. On σ_1 we construct $[n/4]$ points Q_i, where $[n/4]=$"entier" of $n/4$. Further, we construct the corresponding points on the arcs σ_i $(i=2,3,4)$. Together with the vertices of the square we have now $4[n/4]+4$ points. The set $\{\mathscr{P}_i\}_{i=1}^{n}$ is obtained by deletion of $4-n \pmod 4$ extremal points on the arcs σ_i (which are distinct from the vertices of the square). The so obtained set is denoted by P.

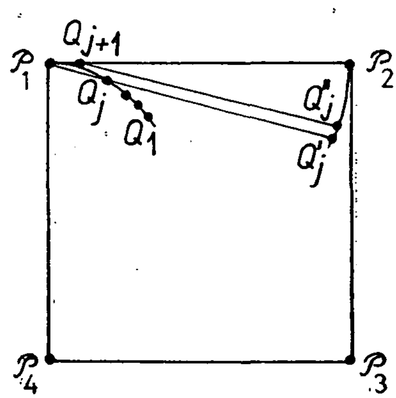


Figure 8. The length of arc $Q_j'Q_j''$ is δ_j

Now we shall verify that all atoms generated by P (equivalently the cells of the corresponding complex on E_2) are triangles, quadrangles or pentagons. For the description of an arbitrary atom α it is sufficent to determine those two subsets of P, which are separated by the lines of the atom. Therefore we shall use the following notations: $\alpha=F|P\setminus F$, where F, $P\setminus F$ are the two subsets in question.

Denote by $\mathcal{M}_i$ the set of points $\{\mathcal{P}_j\}$ lying on σ_i, then $P=\bigcup_{i=1}^{4}\mathcal{M}_i$. By choosing a sufficiently small λ, it is possible to satisfy the following conditions:

(a) Every σ_i is contained in the triangle $\mathcal{P}_i Z \mathcal{P}_j$, where $j=i+1 \pmod 4$, $i\in\{1, 2, 3, 4\}$.

(b) The segments $\mathcal{P}_j Q_1$ $(j=2, 3, 4)$ are intersected by no σ_i $(i=2, 3, 4)$. The same is true for segments, joining $\mathcal{P}_1$ with the endpoints of σ_i $(i=2, 3, 4)$.

(c) The lines g intersect the same σ_i in at most two points.

Let us introduce a classification of the atoms. In our classification we denote by $\mathcal{A}_j$ the classes of the atoms. All the sets F, M in the description of the atoms will be non-empty. Below, the sign $\subset$ denotes only proper inclusion. We put

$$\mathcal{A}_1=\{\emptyset|P\},$$

$$\mathcal{A}_2=\{F|P\setminus F, \text{ where } F\subset\mathcal{M}_i \text{ for some } i\},$$

$$\mathcal{A}_3=\{\mathcal{M}_i|P\setminus\mathcal{M}_i \text{ for some } i\},$$

$$\mathcal{A}_4=\{\mathcal{M}_i\cup\mathcal{M}_j|P\setminus(\mathcal{M}_i\cup\mathcal{M}_j), \text{ where } i\neq j\},$$

$$\mathcal{A}_5=\{\mathcal{M}_i\cup F|P\setminus(\mathcal{M}_i\cup F), \text{ where } F\subset\mathcal{M}_j,\ i\neq j\},$$

$$\mathcal{A}_6=\{F\cup M|P\setminus(F\cup M), \text{ where } F\subset\mathcal{M}_i,\ M\subset\mathcal{M}_j\ (i\neq j) \text{ and every } g \text{ intersects } \sigma_i \text{ and } \sigma_j \text{ in one point only}\},$$

$$\mathcal{A}_7=\{F\cup M|P\setminus(F\cup M), \text{ where } F\subset\mathcal{M}_i,\ M\subset\mathcal{M}_j\ (i\neq j) \text{ and every } g \text{ intersects } \sigma_i \text{ (or } \sigma_j) \text{ in exactly two points}\},$$

$$\mathcal{A}_8=\{\mathcal{M}_i\cup F\cup M|P\setminus(\mathcal{M}_i\cup F\cup M), \text{ where } F\subset\mathcal{M}_k,\ M\subset\mathcal{M}_j \text{ and } i\neq j,\ j\neq k,\ i\neq k.\}$$

By our choice of λ this classification is complete. Direct verification shows that

the atom of $\mathcal{A}_1$ is a quadrangle (by the Lemma),

the atoms of $\mathcal{A}_2$ are either quadrangles, or pentagons (by the construction of $\{Q_i\}$ and (a), (b).),

the atoms of $\mathcal{A}_3$ are either triangles or quadrangles (by the construction of $\{Q_i\}$ and (a), (b)),

the atoms $\mathcal{A}_4$ are either quadrangles or pentagons (by the construction of $\{Q_i\}$ and (a), (b)),

the atoms of $\mathcal{A}_5$ are either quadrangles or pentagons (by the construction of $\{Q_i\}$ and (a), (b)), the atoms of $\mathscr{A}_6$ are quadrangles (by the construction of σ_i and (a), (b)),

the atoms of $\mathscr{A}_7$ are either quadrangles or pentagons (by the construction of $\{Q_i\}$, σ_i and (a), (b), (c)),

the atoms of $\mathscr{A}_8$ are quadrangles (by the construction of σ_i and (a), (b)).

Let us consider one of the types of the atoms, say $\mathscr{A}_2$, in more detail (see Fig. 9). Let g be a line defining an atom from $\mathscr{A}_2$, say $F|P\setminus F$, where $F\subset\mathscr{M}_1$. Then there exist two points Q_k and Q_{k+1} belonging to $\mathscr{M}_1$ such that $Q_k\in P\setminus F$, $Q_{k+1}\in F$. Further let Q'_k and Q'_{k+1} be the points from $\mathscr{M}_4$ corresponding to Q_k and Q_{k+1}. Then the limiting lines are $Q_{k+1}\mathscr{P}_2$, $Q'_{k+1}Q_{k+1}$, Q'_kQ_k, $Q_k\mathscr{P}_2$, $Q'_kQ'_{k+1}$ (by the construction of $\{Q_i\}$ and the choice λ). Hence this atom is a pentagon. If $Q_k=Q_1$ and $n\not\equiv 0 \pmod 4$, then any atom from $\mathscr{A}_2$ is a quadrangle.

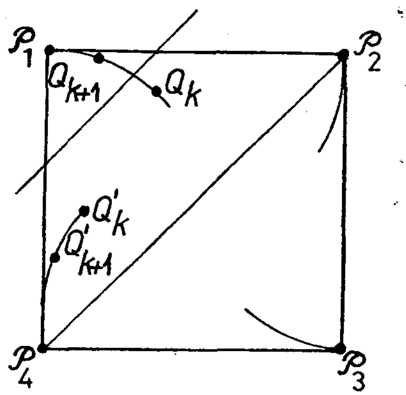


Figure 9

One can similarly treat the other seven cases. This will conclude the proof of the Theorem.

Acknowledgements. The author expresses his gratitude to R. V. Ambartzumian under whose guidance this paper was written.

References

[1] R. V. Ambartzumian, *Combinatorial Integral Geometry: With Applications to Mathematical Stereology*, John Wiley & Sons (Chichester—New York—Brisbane—Toronto—Singapore, 1982).

[2] R. V. Ambartzumian (editor), *Combinatorial Principles in Stochastic Geometry*, Publishing House of the Armenian Academy of Sciences (Yerevan, 1980).

YEREVAN STATE UNIVERSITY
ARMENIA, USSR

Acta Sci. Math., 50 (1986), 67—86

Diagonal lifts of tensor fields to the frame bundle of second order

MANUEL DE LEÓN and MODESTO SALGADO

Introduction

Let M be an n-dimensional manifold of class C^∞, $\mathcal{F}M$ its frame bundle and $\mathcal{F}^2M$ its frame bundle of second order.

The purpose of the present paper is to introduce the (so called) diagonal lifts to $\mathcal{F}^2M$ of tensor fields on M of type $(0, s)$ or $(1, s)$, $s \geqq 1$, with respect to a connection of order 2 on M. A similar theory for $\mathcal{F}M$ and a linear connection on M has been developed by Cordero and one of us in [1]. Actually, the two theories are related.

The paper is structured as follows. In Sections 1 and 2 we recall, for later use, the definitions and properties of the frame bundle of second order $\mathcal{F}^2M$ and of connections of order 2 on M. In Section 3 we introduce a wide class of vector fields on $\mathcal{F}^2M$ and obtain some identities which will be very useful through the rest of the paper. Section 4 is devoted to the definition of the diagonal lift to $\mathcal{F}^2M$ of tensor fields of type $(0, s)$ or $(1, s)$, $s \geqq 1$, with respect to a connection of order 2 on M. The particular cases of the diagonal lifts of tensor fields of type $(1, 1)$ and $(0, 2)$ are the subject of Sections 5 and 6, respectively. We remark that polynomial structures lift into polynomial structures with the same structural polynomial and Riemannian metrics (resp., almost symplectic forms) lift into Riemannian metrics (resp., almost symplectic forms).

Finally, we apply the previous results of Sections 5 and 6 to show that the frame bundle of second order of an almost Hermitian manifold (M, J, G) admits an almost Hermitian structure (J^D, G^D); moreover, the Kaehler form of (J^D, G^D) is the diagonal lift to $\mathcal{F}^2M$ of the Kaehler form of (J, G) and the following result is easily proved: $(\mathcal{F}^2M, J^D, G^D)$ is never an almost Kaehler manifold.

Through the paper, manifolds, tensor fields and connections will be assumed differentiable of class C^∞.

Received March 6, 1984.

1. The frame bundle of order 2

In this section, we recall, for later use the definition and some properties of the frame bundle of order 2. More details can be found in [2], [3], [4], [6] and [7].

Let M be an n-dimensional manifold. If U and V are two neighborhoods of $0 \in R^n$, two mappings $f\colon U \to M$ and $g\colon V \to M$ are said to define the same r-jet at 0 if they have the same partial derivatives up to order r at 0. If f is a diffeomorphism of a neighborhood of 0 onto an open subset of M, then the r-jet $j^r f$ at 0 is called an r-frame at $x = f(0)$. Clearly, an 1-frame is an ordinary linear frame. The set of the r-frames of M, denoted by $\mathscr{F}^r M$, is a principal bundle over M with projection π^r, $\pi^r(j^r f) = f(0)$, and with structure group $G^r(n)$ which will be described next.

Let $G^r(n)$ be the set of r-frames $j^r g$ at $0 \in R^n$, where g is a diffeomorphism from a neighborhood of 0 in R^n onto a neighborhood of 0 in R^n. Then $G^r(n)$ is a Lie group with multiplication defined by the composition of jets, i.e.,

$$(j^r g)(j^r g') = j^r(g \circ g').$$

The group $G^r(n)$ acts of $\mathscr{F}^r M$ on the right by

$$(j^r f)(j^r g) = j^r(f \circ g) \quad \text{for} \quad j^r f \in \mathscr{F}^r M, \quad j^r g \in G^r(n).$$

Clearly, $\mathscr{F}^1 M$ is the bundle of linear frames over M with group $G^1(n) = \mathrm{Gl}(n)$ and projection $\pi^1 = \pi$.

From now on, we shall only consider $\mathscr{F}^1 M$ and $\mathscr{F}^2 M$ and denote $\pi_1^2\colon \mathscr{F}^2 M \to \mathscr{F} M$, $\pi_1^2(j^2 f) = j^1 f$, the canonical projection.

For any coordinate system in M, (U, x^i), we consider the induced coordinate systems $\{\mathscr{F} U, (x^i, X_j^i)\}$ and $\{\mathscr{F}^2 U, (x^i, X_j^i, X_{jk}^i)\}$ in $\mathscr{F} M$ and $\mathscr{F}^2 M$, respectively, where $X_{jk}^i = X_{kj}^i$.

We have a natural isomorphism $G^2(n) \cong \mathrm{Gl}(n) \times S^2(n)$, where $S^2(n)$ is the set of symmetric bilinear forms on R^n, multiplication on the right hand given by $(A, \alpha)(B, \beta) = (AB, \alpha \circ (B, B) + A \circ \beta)$.

Then, the Lie algebra $g^2(n)$ of $G^2(n)$ can be identified to $\mathrm{gl}(n) \oplus S^2(n)$, with a bracket product given by

(1.1)

$$[(A, \alpha), (B, \beta)] = \big([A, B], A \circ \beta - \beta \circ (I, A) - \beta \circ (A, I) - (B \circ \alpha - \alpha \circ (I, B) - \alpha \circ (B, I)\big)$$

where I is the unit matrix.

With these identifications, the adjoint representation of $G^2(n)$ in $a(n) = R^n \oplus \mathrm{gl}(n)$ is given by

$$\mathrm{Ad}^{(2)}(A, \alpha)(v, B) = \big(Av, \bar{\alpha}(v) A^{-1} + ABA^{-1}\big)$$

where $\bar{\alpha}\colon R^n \to \mathrm{gl}(n)$ is the linear map defined by $\bar{\alpha}(v)(w) = \alpha(v, w)$, and the adjoint representation of $G^2(n)$ in $g^2(n)=\mathrm{gl}\,(n)\oplus S^2(n)$ is given by

$$\mathrm{Ad}\,(A,\alpha)(B,\beta)=\big(ABA^{-1},\alpha\circ(A^{-1},BA^{-1})+\alpha\circ(BA^{-1},A^{-1})-$$
$$-ABA^{-1}\circ\alpha\circ(A^{-1},A^{-1})+A\circ\beta\circ(A^{-1},A^{-1})\big).$$

From now on, we shall denote by $\{E_i\}$, $\{E_j^i\}$ and $\{E_{jk}^i\}$, $i,j,k=1,\dots,n$; $E_{jk}^i=E_{kj}^i$; the canonical basis of R^n, $\mathrm{gl}\,(n)$ and $S^2(n)$, respectively.

Since $G^2(n)$ acts on $\mathscr{F}^2M$ on the right, every element (A,α) of the Lie algebra $g^2(n)$ of $G^2(n)$ induces a vector field $\lambda(A,\alpha)$ on $\mathscr{F}^2M$ called the *fundamental vector field corresponding to* (A,α). So, the vertical subspace at any point $p\in\mathscr{F}^2M$ can be decomposed as $\lambda(\mathrm{gl}\,(n))_p\oplus\lambda(S^2(n))_p$.

Let θ be the canonical form on $\mathscr{F}^2M$; θ is an $a(n)$-valued 1-form of type $\mathrm{Ad}^{(2)}(G^2(n))$ and satisfying $\theta(\lambda(A,\alpha))=A$. Let $\theta=\theta_{-1}+\theta_0$ be the decomposition of θ; then, θ_{-1} is an R^n-valued 1-form and θ_0 a $\mathrm{gl}\,(n)$-valued 1-form on $\mathscr{F}^2M$. We have

$$\theta_{-1}(\lambda(A,\alpha))=0,\quad \theta_0(\lambda(A,\alpha))=A.$$

Moreover, $\theta_{-1}=(\pi_1^2)^*\bar\theta$, where $\bar\theta$ is the canonical form of $\mathscr{F}M$. With respect to the canonical bases, we shall put

$$\theta_{-1}=\theta^iE_i,\quad \theta_0=\theta_j^iE_j^i,$$

where θ^i, θ_j^i are locally expressed in $\mathscr{F}^2M$ as

(1.2) $$\theta^i=Y_k^i\,dx^k$$

(1.3) $$\theta_j^i=Y_k^i(dX_j^k-X_{hj}^kY_l^h\,dx^l),$$

(Y_j^i) being the inverse matrix of (X_j^i). From (1.2) and (1.3), we easily obtain the following structure equation

$$d\theta^i=-\theta_k^i\Lambda\theta^k.$$

2. Connections of order 2

A connection Γ in the bundle $\mathscr{F}^2M$ of 2-frames of M is called a connection of order 2 on M.

Let ω be the connection form of Γ; then ω is an 1-form on $\mathscr{F}^2M$ of type $(\mathrm{Ad}\,(G^2(n)))$ and can be decomposed as follows:

(2.1) $$\omega=\omega_0+\omega_1,$$

where ω_0 is a $\mathrm{gl}\,(n)$-valued, and ω_1 an $S^2(n)$-valued 1-form on $\mathscr{F}^2M$. Since $\omega(\lambda(A,\alpha))=(A,\alpha)$, we have

$$\omega_0(\lambda(A,\alpha))=A,\quad \omega_1(\lambda(A,\alpha))=\alpha.$$

Similarly, the curvature form Ω of Γ is a tensorial 2-form on $\mathscr{F}^2 M$ of type Ad $(G^2(n))$ and can be decomposed as

$$\Omega = \Omega_0 + \Omega_1, \tag{2.2}$$

where Ω_0 (resp., Ω_1) is a gl(n)-valued (resp., $S^2(n)$-valued) 2-form on $\mathscr{F}^2 M$. The structure equation is

$$d\omega = -(1/2)[\omega, \omega] + \Omega,$$

or

$$\begin{aligned} d\omega_0 &= -(1/2)[\omega_0, \omega_0] + \Omega_0, \\ d\omega_1 &= -(1/2)\{[\omega_0, \omega_1] + [\omega_1, \omega_0]\} + \Omega_1, \end{aligned} \tag{2.3}$$

taking into account (1.1), (2.1) and (2.2). With respect to the canonical bases, we can write

$$\omega_0 = \omega^i_j E^j_i, \quad \omega_1 = \omega^i_{jk} E^{jk}_i, \quad \Omega_0 = \Omega^i_j E^j_i, \quad \Omega_1 = \Omega^i_{jk} E^{jk}_i,$$

where $\omega^i_{jk} = \omega^i_{kj}$, $\Omega^i_{jk} = \Omega^i_{kj}$, and (2.3) can be equivalently written as

$$d\omega^i_j = -\omega^i_k \Lambda \omega^k_j + \Omega^i_j, \quad d\omega^i_{jk} = -\omega^i_r \Lambda \omega^r_{jk} + \omega^r_k \Lambda \omega^i_{jr} + \omega^r_j \Lambda \omega^i_{rk} + \Omega^i_{jk}.$$

We shall give the following definition.

Definition 2.1. A connection Γ of order 2 on M is said to be partially flat if $\Omega_1 = 0$.

Consequently, a flat connection of order 2 on M is always partially flat.

Let σ be the cross-section of $\mathscr{F}^2 M$ over a coordinate neighborhood (U, x^i) which assigns to each $x \in U$ the 2-frame $(x^i, I, 0)$. We define functions Γ^i_{jk}, Γ^i_{jkl} on U, $\Gamma^i_{jkl} = \Gamma^i_{jlk}$, by

$$\sigma^* \omega_0 = (\Gamma^i_{jk}\, dx^j) E^k_i, \quad \sigma^* \omega_1 = (\Gamma^i_{jkl}\, dx^j) E^{kl}_i.$$

These functions Γ^i_{jk}, Γ^i_{jkl} are called *the components of the connection Γ with respect to the local coordinate system* (U, x^i). By a straightforward computation, we obtain

$$\omega^i_j = Y^i_k (dX^k_j + \Gamma^k_{ml} X^l_j\, dx^m),$$

$$\begin{aligned} \omega^i_{jk} = \{&-\Gamma^t_{ms}(X^s_k Y^c_t Y^i_p X^p_{jc} + X^s_j Y^c_t Y^i_p X^p_{ck} - X^s_{jk} Y^i_t) + \\ &+ \Gamma^t_{msl} X^s_j X^l_k Y^i_t\}\, dx^m - Y^t_r Y^i_s X^s_{tk}\, dX^r_j - Y^t_r Y^i_s X^s_{jt}\, dX^r_k + Y^i_s\, dX^s_{jk}. \end{aligned}$$

Moreover, if we put

$$\sigma^* \Omega_0 = (\sigma^* \Omega^i_j) E^j_i = (1/2)(R^i_{jkl}\, dx^k \Lambda\, dx^l) E^j_i,$$

$$\sigma^* \Omega_1 = (\sigma^* \Omega^i_{jk}) E^{jk}_i = (1/2)(R^i_{jktl}\, dx^t \Lambda\, dx^l) E^{jk}_i,$$

then, from (2.4), we obtain

$$R^i_{jkl} = \partial_k \Gamma^i_{lj} - \partial_l \Gamma^i_{kj} + \Gamma^m_{lj}\Gamma^i_{km} - \Gamma^m_{kj}\Gamma^i_{lm},$$

$$R^i_{jktl} = \partial_t \Gamma^i_{ljk} - \partial_l \Gamma^i_{tjk} + \Gamma^r_{lk}\Gamma^i_{tjr} - \Gamma^r_{tk}\Gamma^i_{ljr} + \Gamma^r_{lj}\Gamma^i_{trk} - \Gamma^r_{tj}\Gamma^i_{lrk} + \Gamma^i_{tr}\Gamma^r_{ljk} - \Gamma^i_{lr}\Gamma^r_{tjk}.$$

We can easily prove that the $n+n^2+n^2(n+1)/2$ global 1-forms θ^i, ω^i_j, ω^i_{jk}, $j \leqq k$, on $\mathscr{F}^2M$ are linearly independent everywhere.

Now, let $p \in \mathscr{F}^2M$; then $(\theta-1)_p$ gives a linear isomorphism of the horizontal subspace H_p at p onto R^n. Thus, we can associate with each $\xi \in R^n$ a horizontal vector field $C(\xi)$ on $\mathscr{F}^2M$ as follows. For each $p \in \mathscr{F}^2M$, $C(\xi)_p$ is the unique horizontal vector at p such that

$$(\theta-1)_p C(\xi)_p = \xi.$$

We call $C(\xi)$ *the standard horizontal vector field on $\mathscr{F}^2M$ corresponding to* ξ.

As a simple computation shows, the local expression of $C(\xi)$ in $\mathscr{F}^2U$ is

$$C(\xi) = X^i_m \xi^m \left\{ \frac{\partial}{\partial x^i} - \Gamma^k_{il} X^l_j \frac{\partial}{\partial X^k_j} - (\Gamma^s_{it} X^t_{jk} + \Gamma^s_{itl} X^t_j X^l_k) \frac{\partial}{\partial X^s_{jk}} \right\}, \tag{2.5}$$

if $\xi = \xi^i E_i$.

Remark that the n global vector fields $C(E_i)$ span the horizontal distribution H_Γ on $\mathscr{F}^2M$. So, the $n+n^2+n^2(n+1)/2$ global vector fields $C(E_i)$, λE^l_j, λE^i_{jk}, $j \leqq k$, define a parallelism on $\mathscr{F}^2M$ and are dual to σ^i, ω^i_j, ω^i_{jk}; moreover the local expressions of λE^i_j, λE^i_{jk} on $\mathscr{F}^2U$ are

$$\lambda E^i_j = X^t_i \frac{\partial}{\partial X^t_j} + X^t_{is} \frac{\partial}{\partial X^t_{js}} + X^t_{si} \frac{\partial}{\partial X^t_{sj}}, \tag{2.6}$$

$$\lambda E^i_{jk} = X^t_i \frac{\partial}{\partial X^t_{jk}}. \tag{2.7}$$

From (2.5), (2.6) and (2.7), we notice that the horizontal distribution H_Γ is spanned by the local vector fields

$$D_i = \frac{\partial}{\partial x^i} - \Gamma^k_{il} X^l_r \frac{\partial}{\partial X^k_r} - \{\Gamma^s_{it} X^t_{rk} + \Gamma^s_{itl} X^t_r X^l_k\} \frac{\partial}{\partial X^s_{rk}}$$

and the vertical distribution V is spanned by the local vector fields

$$D^i_j = \frac{\partial}{\partial X^i_j} + Y^r_j X^t_{rs} \frac{\partial}{\partial X^t_{is}} + Y^r_j X^t_{sr} \frac{\partial}{\partial X^t_{si}}, \quad D^i_{jk} = \frac{\partial}{\partial X^i_{jk}}$$

when we restrict ourselves to $\mathscr{F}^2U$.

The frame $\{D_i, D^i_j, D^i_{jk}\}$ is adapted to the almost product structure $(H_\Gamma; V)$ and we call it the *adapted frame on $\mathscr{F}^2U$*. The local 1-forms η^i, η^i_j, η^i_{jk} on $\mathscr{F}^2U$ dual to $\{D^i, D^i_j, D^i_{jk}\}$ are given by

$$\eta^i = dx^i, \quad \eta^i_j = \Gamma^i_{rs} X^s_j \, dx^r + dX^i_j,$$
$$\eta^i_{jk} = \{(\Gamma^i_{tr} X^r_{jk} + \Gamma^i_{trs} X^r_j X^s_k) - Y^s_m \Gamma^m_{tr} (X^r_j X^i_{sk} + X^r_k X^i_{js})\} \, dx^t -$$
$$- Y^r_t (\delta^{sj} X^i_{rk} + \delta^{sk} X^i_{jr}) \, dX^t_s + dX^i_{jk},$$

and $\{\eta^i, \eta^i_j, \eta^i_{jk}\}$ will be called the *adapted coframe on* $\mathscr{F}^2 U$.

Now, let Γ be a connection of order 2 on M. Since the canonical projection π^2_1: $\mathscr{F}^2 M \to \mathscr{F} M$ is a homomorphism of principal bundles over the identity of M inducing the canonical projection $G^2(n) \to \mathrm{Gl}\,(n)$, then the connection Γ defines a connection in $\mathscr{F} M$, that is, a linear connection $\bar{\Gamma}$ on M. We call $\bar{\Gamma}$ the *linear connection on* M *induced from* Γ. If $\bar{\omega}$, $\bar{\Omega}$ are the connection and the curvature forms of $\bar{\Gamma}$, then

$$(2.8) \qquad (\pi^2_1)^* \bar{\omega} = \omega_0, \quad (\pi^2_1)^* \bar{\Omega} = \Omega_0.$$

Let $\bar{\lambda} A$ (resp., $B(\xi)$) be the fundamental vector field (resp., the standard horizontal vector field with respect to $\bar{\Gamma}$) corresponding to $A \in \mathrm{gl}\,(n)$ (resp., $\xi \in R^n$). A simple computation shows that

$$(2.9) \qquad \pi^2_1 \lambda(A, \alpha) = \bar{\lambda} A \quad (\text{resp.}, \ \pi^2_1 C(\xi) = B(\xi)).$$

If $\bar{\sigma}$ is the cross section of $\mathscr{F} M$ over a coordinate neighborhood (U, x^i) which assigns to each $x \in U$ the linear frame (x^i, I), then we can define the components of $\bar{\Gamma}$ by

$$\sigma^* \bar{\omega} = (\bar{\Gamma}^i_{jk} \, dx^j) E^k_i.$$

Taking into account (2.8), we find $\bar{\Gamma}^i_{jk} = \Gamma^i_{jk}$ and, consequently, $\bar{R}^i_{jkl} = R^i_{jkl}$, $\bar{R}^i_{jkl}$ being the components of the tensor curvature of $\bar{\Gamma}$.

3. Vector fields on $\mathscr{F}^2 M$

Our aim in this section is to introduce a wide class of vector fields on $\mathscr{F}^2 M$ and obtain some identities which will be very useful through the rest of the paper. Previously, we shall consider an arbitrary principal bundle $P(M, G)$ over M with structure group G. Let $\mathfrak{g}$ be the Lie algebra of G; then, for any function f: $P \to \mathfrak{g}$, we can define the vertical vector field λf by

$$(3.1) \qquad (\lambda f)(p) = (\lambda(f(p)))_p, \quad p \in P.$$

It will be called the *fundamental vector field corresponding to* f. On the other hand, if Γ is a connection in P, we shall denote by X^H the horizontal lift to P of a vector field X on M. Then, we have

Proposition 3.1. *Let Γ be a connection in P with connection form ω and curvature form Ω. Then*

$$(1)\qquad [\lambda f, \lambda g] = \lambda[f, g] + \lambda(\lambda f(g)) - \lambda(\lambda g(f)),$$

$$(2)\qquad [X^H, \lambda g] = \lambda(X^H g),$$

$$(3)\qquad [X^H, Y^H] = [X, Y]^H - 2\lambda\Omega(X^H, Y^H),$$

for any vector fields X, Y on M and any functions $f, g\colon P \to \mathfrak{g}$.

Proof. Let us recall the structure equation of Γ:

$$d\omega = -(1/2)[\omega, \omega] + \Omega.$$

Now, from (3.1), we have $\omega(\lambda f) = f$, $\omega(\lambda g) = g$ and, therefore,

$$2d\omega(\lambda f, \lambda g) = \lambda f(\omega(\lambda g)) - \lambda g(\omega(\lambda f)) - \omega([\lambda f, \lambda g]) =$$
$$= \lambda f(g) - \lambda g(f) - \omega([\lambda f, \lambda g]).$$

On the other hand, since $\Omega(\lambda f, \lambda g) = 0$, we obtain

$$[\omega, \omega](\lambda f, \lambda g) = [f, g].$$

Hence

$$\omega([\lambda f, \lambda g]) = [f, g] + (\lambda f)g - (\lambda g)f$$

and, taking into account that $[\lambda f, \lambda g]$ is vertical, we deduce (1).

(2) follows by a similar device, considering that $[X^H, \lambda f]$ is vertical.

To prove (3), it suffices to recall that the horizontal component of $[X^H, Y^H]$ is $[X, Y]^H$ and, as a direct consequence of the structure equation,

$$\omega([X^H, Y^H]) = -2\Omega(X^H, Y^H).$$

Then,

$$[X^H, Y^H] = [X, Y]^H - 2\lambda\Omega(X^H, Y^H).$$

We remark that if f and g are the constant functions A and B, respectively, $A, B \in \mathfrak{g}$, then the Proposition 3.1 gives the well-known formulas

$$(3.2)\qquad [\lambda A, \lambda B] = \lambda[A, B], \quad [X^H, \lambda B] = 0.$$

From now on, we return to the frame bundle of second order $\mathscr{F}^2 M$ of a manifold M with a connection Γ of order 2. If X is a vector field on M with local expressions $X = X^i(\partial/\partial x^i)$ in a coordinate neighborhood U in M, then the local expression of X^H in $\mathscr{F}^2 U$ with respect to the induced coordinates can be obtained by a direct computation and it is given by

$$(3.3)\qquad X^H = X^i\left\{\frac{\partial}{\partial x^i} - \Gamma^r_{il} X^l_j \frac{\partial}{\partial X^r_j} - (\Gamma^r_{il} X^l_{jk} + \Gamma^r_{ilm} X^l_j X^m_k)\frac{\partial}{\partial X^r_{jk}}\right\}$$

or

$$(3.4)\qquad X^H = X^i D_i,$$

with respect to the adapted frame.

From (2.9) and (3.3), one easily deduces that $\pi_1^2 X^H = X^H$, X^H being the horizontal lift of X to $\mathscr{F}M$ with respect to the induced connection $\bar{\Gamma}$.

Now, let F be a tensor field of type (1, 1) on M. We can define a function F°: $\mathscr{F}M \to \mathrm{gl}(n)$ as follows: For any $p \in \mathscr{F}M$, $F^\circ(p)$ is the matrix representation of F_x with respect to p, $x = \pi(p)$. The function F induces a function on $\mathscr{F}^2 M$, also denoted by F°, by putting $F^\circ = F^\circ \circ \pi_1^2$. If F has local components F_i^h in U, then we have in $\mathscr{F}^2 U$

$$F^\circ = [F_t^h X_j^t Y_h^i],$$

and the corresponding fundamental vector field is given by

$$\lambda F^\circ = F_t^r X_j^t \frac{\partial}{\partial X_j^r} + F_t^h X_j^t Y_h^i X_{is}^r \frac{\partial}{\partial X_{js}^r} + F_t^h X_j^t Y_h^i X_{si}^r \frac{\partial}{\partial X_{sj}^r}. \tag{3.5}$$

Moreover, if $A \in \mathrm{gl}(n)$ and $\alpha \in S^2(n)$, we can consider the functions

$$F^\circ A\colon \mathscr{F}^2 M \to \mathrm{gl}(n), \quad F^\circ \alpha\colon \mathscr{F}^2 M \to S^2(n)$$

defined by

$$(F^\circ A)(p) = F^\circ(p) A, \quad (F^\circ \alpha)(p) = F^\circ(p) \circ \alpha, \quad p \in \mathscr{F}^2 M.$$

The corresponding fundamental vector fields $\lambda(F^\circ A)$ and $\lambda(F^\circ \alpha)$ are locally expressed in $\mathscr{F}^2 U$ as

$$\lambda(F^\circ A) = F_j^h X_s^j A_t^s \frac{\partial}{\partial X_t^h} + F_j^h X_s^j Y_h^r A_t^s X_{rk}^i \frac{\partial}{\partial X_{tk}^i} + F_j^h X_s^j Y_h^r A_t^s X_{kr}^i \frac{\partial}{\partial X_{kt}^i}, \tag{3.6}$$

$$\lambda(F^\circ \alpha) = F_j^h X_s^j \alpha_{mn}^s \frac{\partial}{\partial X_{mn}^h}; \tag{3.7}$$

where $A = A_j^i E_i^j$, $\alpha = \alpha_{jk}^i E_i^{jk}$, $\alpha_{jk}^i = \alpha_{kj}^i$.

The following formulas will be useful and can be obtained by a straightforward computation taking into account (3.3), (3.5), (3.6), (3.7) and Proposition 3.1:

$$\begin{aligned}
&[X^H, \lambda F^\circ] = \lambda(\nabla_X F)^\circ, \\
&[X^H, \lambda(F^\circ A)] = \lambda((\nabla_X F)^\circ A), \\
&[X^H, \lambda(F^\circ \alpha)] = \lambda((\nabla_X F)^\circ \alpha), \\
&[\lambda(F^\circ A), \lambda B] = \lambda(F^\circ [A, B]), \\
&[\lambda(F^\circ A), \lambda \alpha] = \lambda([F^\circ A, \alpha]) = \lambda\{F^\circ(A \circ \alpha) - \alpha \circ (I, F^\circ A) - \alpha \circ (F^\circ A, I)\}, \\
&[\lambda(F^\circ A), \lambda(F^\circ B)] = \lambda((F^2)^\circ [A, B]), \\
&[\lambda(F^\circ A), \lambda(F^\circ \beta)] = \lambda\{(F^2)^\circ (A \circ \beta) - (F^\circ \beta) \circ (I, F^\circ A) - (F^\circ \beta) \circ (F^\circ A, I)\}, \\
&[\lambda(F^\circ \alpha), \lambda B] = \lambda(F^\circ [\alpha, B]) = -\lambda\{F^\circ(B \circ \alpha) - F^\circ(\alpha \circ (I, B)) - F^\circ(\alpha \circ (B, I))\}, \\
&[\lambda(F^\circ \alpha), \lambda \beta] = 0, \\
&[\lambda(F^\circ \alpha), \lambda(F^\circ \beta)] = 0,
\end{aligned} \tag{3.8}$$

for any vector field X, any tensor field F of type (1, 1) on M, any $A, B \in \mathrm{gl}(n)$, $\alpha, \beta \in S^2(n)$, where ∇ denotes the covariant derivative of the linear connection $\bar{\Gamma}$ induced from Γ.

Moreover, (1) in Proposition (3.1) can be equivalently written as follows:

$$[X^H, Y^H] = [X, Y]^H - 2\lambda\Omega_0(X^H, Y^H) - 2\lambda\Omega_1(X^H, Y^H).$$

Now, a simple computation shows that

$$2\lambda\Omega_0(X^H, Y^H) = \lambda(R(X, Y)^\circ),$$

R being the curvature tensor of $\bar{\Gamma}$. Then, we have

$$(3.9) \qquad [X^H, Y^H] = [X, Y]^H - \lambda(R(X, Y)^\circ) - 2\lambda\Omega_1(X^H, Y^H),$$

for any vector fields X, Y on M.

4. Diagonal lifts of tensor fields

Let $u \in (R^n)^*$; then, there exist $u' \in \mathrm{gl}(n)^*$ and $u'' \in S^2(n)^*$ canonically associated to u and given by

$$u'(A) = \sum_{j=1}^{n} u(A_j), \quad A \in \mathrm{gl}(n), \quad \text{and} \quad u''(\alpha) = \sum_{j,k=1}^{n} u(\alpha_{jk}), \quad \alpha \in S^2(n),$$

where A_j (resp., α_{jk}) denotes the j^{th} column (resp., the (j, k)-column) of A (resp., α).

Let $u \in \mathrm{Hom}(R^n, R^n)$; then there exist

$$u' \in \mathrm{Hom}(\mathrm{gl}(n), \mathrm{gl}(n)) \quad \text{and} \quad u' \in \mathrm{Hom}(S^2(n), S^2(n))$$

canonically associated to u and given by

$$u'(A) = u \circ A, \quad A \in \mathrm{gl}(n), \quad \text{and} \quad u''(\alpha) = u \circ a, \quad \alpha \in S^2(n).$$

It is easy to show that if $\mathrm{rank}\, u = r$, then $\mathrm{rank}\, u' = rn$ and $\mathrm{rank}\, u'' = rn(n+1)/2$.

These two definitions of u' and u'' can be extended as follows: Let $u \in \otimes_s(R^n)^*$, $s \geqq 1$; then there exist $u' \in \otimes_s \mathrm{gl}(n)^*$ and $u'' \in \otimes_s S^2(n)^*$ canonically associated to u and given by

$$u'(A_1, \ldots, A_s) = \sum_{j=1}^{n} u((A_1)_j, \ldots, (A_s)_j), \quad A_1, \ldots, A_s \in \mathrm{gl}(n),$$

and

$$u''(\alpha_1, \ldots, \alpha_s) = \sum_{j,k=1}^{n} u((\alpha_1)_{jk}, \ldots, (\alpha_s)_{jk}), \quad \alpha_1, \ldots, \alpha_s \in S^2(n),$$

where $(A_i)_j$ (resp., $(\alpha_i)_{jk}$) denotes the j^{th} column (resp., the (j, k)-column) of A_i (resp., α_i).

In particular, if $s=2$, we have

$$u'(A, B) = \sum_{j=1}^{n} u(A_j, B_j), \quad A, B \in \mathrm{gl}(n),$$

$$u''(\alpha, \beta) = \sum_{j,k=1}^{n} u(\alpha_{jk}, \beta_{jk}), \quad \alpha, \beta \in S^2(n),$$

and it is easy to show that if u is symmetric (resp., skewsymmetric), then u' and u'' are also symmetric (resp., skewsymmetric). Moreover, if rank $u=r$, then rank $u'=rn$ and rank $u''=rn(n+1)/2$.

Now, let $u \in R^n \otimes (\otimes_s (R^n)^*)$, $s \geqq 1$; then there exist $u' \in \mathrm{gl}(n) \otimes (\otimes_s \mathrm{gl}(n)^*)$ and $u'' \in S^2(n) \otimes (\otimes_s S^2(n)^*)$ canonically associated to u such that the j^{th} column (resp., the (j, k)-column) of $u'(A_1, \dots, A_s)$ (resp., $u'(\alpha_1, \dots, \alpha_s)$) is $u((A_1)_j, \cdots, (A_s)_j)$ (resp., $u((\alpha_1)_{jk}, \dots, (\alpha_s)_{jk})$).

Let τ be an 1-form on M. For each $p \in \mathscr{F}^2 M$, we put

$$u_p = \tau_x \circ \pi_1^2(p), \tag{4.1}$$

where $\pi_1^2(p)\colon R^n \to T_x M$, $x=\pi^2(p)$, is considered as a linear isomorphism. Then

Definition 4.1. The diagonal lift τ^D of τ to $\mathscr{F}^2 M$ is the 1-form given by

$$(\tau^D)_p(X) = u_p((\theta_{-1})_p X) + u'_p((\omega_0)_p X) + u''_p((\omega_1)_p X),$$

$X \in T_p(\mathscr{F}^2 M)$, $p \in \mathscr{F}^2 M$, where $u'_p \in \mathrm{gl}(n)^*$, $u''_p \in S^2(n)^*$ are the elements associated to $u_p \in (R^n)^*$ given by (4.1).

If $\tau = \tau_i\, dx^i$ is the local expression of τ in a coordinate neighborhood U, then the local expression of τ^D in $\mathscr{F}^2 U$ with respect to the adapted frame field is

$$\tau^D = \tau_i \eta^i + \sum_{j=1}^{n} \tau_i \eta^i_j + \sum_{j,k=1}^{n} \tau_i \eta^i_{jk}.$$

The definition above can be extended to an arbitrary covariant tensor field as follows. Let G be a tensor field on M of type $(0, s)$, $s \geqq 1$; for each $p \in \mathscr{F}^2 M$, we put

$$u_p = G_x \circ (\pi_1^2(p) \times \overset{s}{\dots} \times \pi_1^2(p)), \quad x = \pi_1^2(p). \tag{4.2}$$

Definition 4.2. The diagonal lift G^D of G to $\mathscr{F}^2 M$ is the tensor field of the same type given by

$$G^D(X_1, \dots, X_s) = u_p((\theta_{-1})_p X_1, \dots, (\theta_{-1})_p X_s) + u'_p((\omega_0)_p X_1, \dots, (\omega_0)_p X_s) + \\ + u''_p((\omega_1)_p X_1, \dots, (\omega_1)_p (X_s),$$

$X_1, \dots, X_s \in T_p(\mathscr{F}^2 M)$, $p \in \mathscr{F}^2 M$, where $u'_p \in \otimes_s \mathrm{gl}(n)^*$, $u''_p \in \otimes_s S^2(n)^*$ are the elements associated to $u_p \in \otimes_s (R^n)^*$, given by (4.2).

If $G_{j_1\dots j_s}$ are the local components of G in U, then

$$G^D = G_{j_1\dots j_s}\eta^{j_1}\otimes\dots\otimes\eta^{j_s}+\sum_l \delta_l^{i_1}\dots\delta_l^{i_s}G_{j_1\dots j_s}\eta_{i_1}^{j_1}\otimes\dots\otimes\eta_{i_s}^{j_s}+ \tag{4.3}$$

$$+\sum_{l,m}\delta_l^{i_1}\delta_m^{j_1}\dots\delta_l^{i_s}\delta_m^{j_s}G_{k_1\dots k_s}\eta_{i_1j_1}^{k_1}\otimes\dots\otimes\eta_{i_sj_s}^{k_s}$$

is the local expression of G^D with respect to the adapted frame field.

Now, let F be an arbitrary tensor field of type (1, 1) on M. For each $p\in\mathcal{F}^2M$, we put

$$u_p = (\pi_1^2(p))^{-1}\circ F_x\circ\pi_1^2(p),\quad x=\pi^2(p). \tag{4.4}$$

Then

Definition 4.3. The diagonal lift F^D of F to $\mathcal{F}^2M$ is the tensor field of type (1, 1) given by

$$(F^D)_pX = C(u_p((\theta_{-1})_pX))+\lambda(u'_p((\omega_0)_pX))+\lambda(u''_p((\omega_1)_pX)),$$

$X\in T_p(\mathcal{F}^2M)$, $p\in\mathcal{F}^2M$, where $u'_p\in\mathrm{Hom}(\mathrm{gl}(n), \mathrm{gl}(n))$, $u''_p\in\mathrm{Hom}(S^2(n), S^2(n))$ are the elements associated to $u_p\in\mathrm{Hom}(R^n, R^n)$ given by (4.4).

If F_j^h are the local components of F in U, then

$$F^D = F_j^hD_h\otimes\eta^j+\delta_j^iF_k^hD_h^i\otimes\eta_j^k+\delta_j^i\delta_l^kF_m^hD_{jl}^h\otimes\eta_{ik}^m \tag{4.5}$$

is the local expression of F^D with respect to the adapted frame field in $\mathcal{F}^2U$.

Definition 4.3 can be extended as follows: Let F be a tensor field of type $(1, s)$, $s\geqq 1$; for each $p\in\mathcal{F}^2U$, we put

$$u_p = (\pi_1^2(p))^{-1}\circ F_x\circ(\pi_1^2(p)\times\overset{s}{\dots}\times\pi_1^2(p)),\quad x=\pi^2(p). \tag{4.6}$$

Then

Definition 4.4. The diagonal lift F^D of F to $\mathcal{F}^2M$ is the tensor field of type $(1, s)$ given by

$$(F^D)_p(X_1, \dots, X_s) = C(u_p((\theta_{-1})_pX_1, \dots, (\theta_{-1})_pX_s))+$$
$$+\lambda(u'_p((\omega_0)_pX_1, \dots, (\omega_0)_pX_s))+\lambda(u''_p((\omega_1)_pX_1, \dots, (\omega_1)_pX_s));$$

$X_1, \dots, X_s\in T_p(\mathcal{F}^2M)$, $p\in\mathcal{F}^2M$, where $u'_p\in\mathrm{gl}(n)\otimes(\otimes_s\mathrm{gl}(n)^*)$, $u''_p\in S^2(n)\otimes(\otimes_sS^2(n)^*)$ are the elements associated to $u_p\in R^n\otimes(\otimes_s(R^n)^*)$ given by (4.6).

If $F_{j_1\dots j_s}^h$ are the local components of F in U, then

$$F^D = F_{j_1\dots j_s}^hD_h\otimes\eta^{j_1}\otimes\dots\otimes\eta^{j_s}+\sum_{l,h}\delta_l^{i_1}\dots\delta_l^{i_s}F_{j_1\dots j_s}^hD_h^l\otimes\eta_{i_1}^{j_1}\otimes\dots\otimes\eta_{i_s}^{j_s}+$$
$$+\sum_{l,m,h}\delta_l^{i_1}\delta_m^{j_1}\dots\delta_l^{i_s}\delta_m^{j_s}F_{k_1\dots k_s}^hD_{lm}^h\otimes\eta_{i_1j_1}^{k_1}\otimes\dots\otimes\eta_{i_sj_s}^{k_s}$$

is the local expression of F^D in $\mathcal{F}^2U$ with respect to the adapted frame field.

Now, let $\bar{\Gamma}$ be the linear connection on M induced from a connection Γ of order 2. The diagonal lifts to $\mathscr{F}M$ of tensor fields on M with respect to a linear connection on M have been considered in [1] by CORDERO and DE LEÓN. Actually, we can easily prove that the diagonal lifts of tensor fields to $\mathscr{F}^2M$ with Γ projects canonically to the diagonal lifts to $\mathscr{F}M$ with respect to $\bar{\Gamma}$.

5. Diagonal lifts of tensor fields of type (1, 1)

We shall now study the diagonal lift F^D of a tensor field F of type (1, 1) in more detail.

Proposition 5.1. *We have*

(1) $F^D X^H = (FX)^H$,

(2) $F^D(\lambda f) = \lambda(F^\circ f)$,

(3) $F^D(\lambda g) = \lambda(F^\circ g)$,

(4) $F^D(\lambda A) = \lambda(F^\circ A)$,

(5) $F^D(\lambda\alpha) = \lambda(F^\circ\alpha)$,

for any vector field X on M, any function $f\colon \mathscr{F}^2M \to \mathrm{gl}(n)$, any function $g\colon \mathscr{F}^2M \to S^2(n)$, any $A \in \mathrm{gl}(n)$ and any $\alpha \in S^2(n)$.

Proof. (1), (2) and (3) follow directly from (3.1), (3.4) and Definition 4.3 taking into account that $F^\circ(p) = u_p$; $p \in \mathscr{F}^2M$. (4) (resp., (5)) is a direct consequence of (2) (resp., (3)), when one considers the constant function A (resp., α).

From Proposition 5.1, we obtain

Proposition 5.2. *Let F, G be tensor fields of type* (1, 1) *on M and denote by I the identity tensor field. Then*

(1) $(FG)^D = F^D G^D$,

(2) $I^D = I$.

Proof. To prove (1) it suffices to check the identities

$$(FG)^D(X^H) = F^D(G^D(X^H)); \quad (FG)^D(\lambda A) = F^D(G^D(\lambda A)); \quad (FG)^D(\lambda\alpha) = F^D(G^D(\lambda\alpha));$$

for any vector field X on M, any $A \in \mathrm{gl}(n)$ and any $\alpha \in S^2(n)$. The first one follows from (1) in Proposition 5.1 and the other identities follow from (2), (3), (4) and (5) in Proposition 5.1 taking into account that $(FG)^\circ = F^\circ G^\circ$. On the other hand, we have

$$I^D X^H = (IX)^H = X^H, \quad I^D(\lambda A) = \lambda(I^\circ A) = \lambda A, \quad I^D(\lambda\alpha) = \lambda(I^\circ \alpha) = \lambda\alpha,$$

because I° is the constant function I. Thus, Proposition 5.2 is proved.

As a direct consequence of Proposition 5.2, we have

Proposition 5.3. *If $P(t)$ is a polynomial in one variable t, then*

$$(P(F))^D = P(F^D).$$

Corollary 5.4. *Let F be a tensor field of type* (1, 1) *on M. Then, if F defines on M a polynomial structure of* rank r *and structural polynomial $P(t)=0$, its diagonal lift F^D defines on $\mathscr{F}^2M$ a polynomial structure of* rank $r(1+n+n(n+1)/2)$ *and with the same structural polynomial. In particular, if F is an almost complex structure (resp., an f-structure of* rank r*) on M, then F^D is an almost complex structure (resp., an f-structure of* rank $r(1+n+n(n+1)/2)$*) on $\mathscr{F}^2M$.*

Denote by N_{F^D} and N_F the Nijenhuis tensor of F^D and F, respectively. Thus, taking into account the definition of the Nijenhuis tensor, the formulas (3.8) and (3.9) and Proposition 5.1, we find by a straightforward computation the following identities:

$$N_{F^D}(X^H, Y^H) = (N_F(X, Y))^H - \lambda((R(FX, FY) - FR(FX, Y) - FR(X, FY) +$$
$$+ F^2R(X, Y))^\circ) - 2\lambda(\Omega_1((FX)^H, (FY)^H) - F^\circ\Omega_1((FX)^H, Y^H) -$$
$$- F^\circ\Omega_1(X^H, (FY)^H) + (F^2)^\circ\Omega_1(X^H, Y^H)),$$
$$N_{F^D}(X^H, \lambda B) = \lambda((\nabla_{FX}F - F\nabla_X F)^\circ B), \quad N_{F^D}(X^H, \lambda\beta) = \lambda((\nabla_{FX}F - F\nabla_X F)^\circ \beta),$$
$$N_{F^D}(\lambda A, \lambda B) = N_{F^D}(\lambda A, \lambda\beta) = N_{F^D}(\lambda\alpha, \lambda\beta) = 0,$$

for any vector fields X, Y on M, any $A, B \in \mathrm{gl}\,(n)$ and any $\alpha, \beta \in S^2(n)$. Therefore, we have

Proposition 5.5. *Let F be a tensor field on M of type* (1, 1) *and F^D its diagonal lift to $\mathscr{F}^2M$. Then the condition $N_{F^D}=0$ is equivalent to the conditions*

$$N_F = 0, \quad F\nabla_X F - \nabla_{FX}F = 0,$$
$$R(FX, FY) - FR(FX, Y) - FR(X, FY) + F^2R(X, Y) = 0,$$
$$\Omega_1((FX)^H, FY)^H) - F^\circ\Omega_1((FX)^H, Y^H) - F^\circ\Omega_1(X^H, (FY)^H) + (F^2)^\circ\Omega_1(X^H, Y^H) = 0,$$

for arbitrary vector fields X, Y on M. The three last conditions can be equivalently *written as*

$$F_i^k \nabla_k F_j^h - F_k^h \nabla_i F_j^h = 0,$$

$$R_{klm}^h F_j^l F_i^m - R_{kli}^m F_j^l F_m^h - R_{kjl}^m F_i^l F_m^h + R_{kji}^l F_l^m F_m^h = 0,$$

$$R_{kjlm}^h F_s^l F_t^m - R_{kjlt}^i F_s^l F_i^h - R_{kjsm}^i F_i^h F_t^m + R_{kjst}^i F_r^h F_i^r = 0,$$

where F_j^h are the local components of F.

To obtain some meaningful formulas on Lie derivatives, let us recall that the Lie derivative $\mathscr{L}_{\tilde{X}}\tilde{F}$ of a tensor field $\tilde{F}$ of type $(1, 1)$ with respect to a vector field $\tilde{X}$ is defined by

$$(\mathscr{L}_{\tilde{X}}\tilde{F})(\tilde{Y}) = [\tilde{X}, \tilde{F}\tilde{Y}] - \tilde{F}[\tilde{X}, \tilde{Y}],$$

$\tilde{Y}$ being an arbitrary vector field. Thus, taking into account (3.8), (3.9) and Proposition 5.1, we obtain

$$(\mathscr{L}_{X^H} F^D)(Y^H) = ((\mathscr{L}_X F)(Y))^H - \lambda((R(X, FY) - FR(X, Y))^\circ) -$$
$$-2\lambda(\Omega_1(X^H, (FY)^H) - F^\circ \Omega_1(X^H, Y^H)),$$
$$(\mathscr{L}_{X^H} F^D)(\lambda A) = \lambda((\nabla_X F)^\circ A), \quad (\mathscr{L}_{X^H} F^D)(\lambda\alpha) = ((\nabla_X F)^\circ \alpha),$$

for any vector fields X, Y on M, any $A \in \mathrm{gl}(n)$ and any $\alpha \in S^2(n)$. Thus, we have

Proposition 5.6. *Let X be a vector field and F a tensor field of type* $(1, 1)$ *on M. Then the condition $\mathscr{L}_{X^H} F^D = 0$ is equivalent to the conditions*

$$\mathscr{L}_X F = 0, \quad \nabla_X F = 0, \quad R(X, FY) - FR(X, Y) = 0,$$
$$\Omega_1(X^H, (FY)^H) - F^\circ \Omega_1(X^H, Y^H) = 0,$$

for any vector field Y on M. The two last conditions can be equivalently written as

$$X^j (R_{kjl}^h F_i^l - R_{kji}^l F_l^h) = 0, \quad X^m (R_{kjlm}^h F_i^l - R_{kjim}^l F_l^h) = 0,$$

where X^j and F_j^h are the local components of X and F, respectively.

If we next take into account (3.7) and Proposition 5.1, we find

$$(\mathscr{L}_{\lambda A} F^D)(Y^H) = (\mathscr{L}_{\lambda A} F^D)(\lambda B) = (\mathscr{L}_{\lambda A} F^D)(\lambda\beta) = 0,$$
$$(\mathscr{L}_{\lambda\alpha} F^D)(Y^H) = (\mathscr{L}_{\lambda\alpha} F^D)(\lambda B) = (\mathscr{L}_{\lambda\alpha} F^D)(\lambda\beta) = 0,$$

for any vector field Y on M, any $A, B \in \mathrm{gl}(n)$ and any $\alpha, \beta \in S^2(n)$. Thus, we have

Proposition 5.7. *Let F be a tensor field on M of type* $(1, 1)$. *Then* $\mathscr{L}_{\lambda A} F^D = \mathscr{L}_{\lambda\alpha} F^D = 0$, *for any* $A \in \mathrm{gl}(n)$ *and any* $\alpha \in S^2(n)$.

6. Diagonal lifts of tensor fields of type (0, 2)

Let G be a tensor field of type $(0, 2)$ on M. Particularizing in this case Definition 4.2 and (4.3), we have that the diagonal lift G^D of G to $\mathscr{F}^2 M$ is a tensor field of type $(0, 2)$ on $\mathscr{F}^2 M$ with local expression in $\mathscr{F}^2 U$

$$G^D = G_{ij}\eta^i \otimes \eta^j + \delta^{kl} G_{ij}\eta^i_k \otimes \eta^j_l + \delta^{km}\delta^{ln} G_{ij}\eta^i_{kl} \otimes \eta^j_{mn}, \tag{6.1}$$

G_{ij} being the local components of G in U.

By using (6.1), one easily deduces that if G has constant rank r, then G^D has constant rank $r(1+n+n(n+1)/2)$. Thus, we have

Proposition 6.1. (1) *If G is a Riemannian metric on M, then G^D is a Riemannian metric on $\mathscr{F}^2 M$.*

(2) *If G is an almost symplectic form on M, then G^D is an almost symplectic form on $\mathscr{F}^2 M$.*

Let us now introduce two new definitions: Let $A, B \in \mathrm{gl}(n)$, $\alpha, \beta \in S^2(n)$ and let G be a tensor field of type $(0, 2)$ on M with local components G_{ij}; then

$$G^\circ(A, B) = \delta^{rs} A^k_r B^l_s X^i_k X^j_l G_{ij} \tag{6.2}$$

and

$$G^\circ(\alpha, \beta) = \delta^{rt}\delta^{sn}\alpha^k_{rs}\beta^l_{tn} X^i_k X^j_l G_{ij} \tag{6.3}$$

are globally well-defined functions on $\mathscr{F}^2 M$, where

$$A = A^i_j E^j_i, \quad B = B^i_j E^j_i, \quad \alpha = \alpha^i_{jk} E^{jk}_i, \quad \beta = \beta^i_{jk} E^{jk}_i.$$

The following formulas are easily obtained:

$$\begin{gathered} G^D(\lambda A, \lambda B) = G^\circ(A, B), \quad G^D(\lambda A, \lambda\beta) = G^D(\lambda\beta, \lambda A) = 0, \\ G^D(\lambda A, X^H) = G^D(X^H, \lambda A) = G^D(\lambda\alpha, X^H) = G^D(X^H, \lambda\alpha) = 0, \\ G^D(\lambda\alpha, \lambda\beta) = G^\circ(\alpha, \beta), \quad G^D(X^H, Y^H) = (G(X, Y))^V, \end{gathered} \tag{6.4}$$

for any $A, B \in \mathrm{gl}(n)$, $\alpha, \beta \in S^2(n)$ and arbitrary vector fields X, Y on M, where $f^V = f \circ \pi^2$, for any function f on M.

Next, we shall compute the Lie derivatives of G^D with respect to vector fields λA, $\lambda\alpha$ or X^H on $\mathscr{F}^2 M$. To do this, let us recall that the Lie derivative $\mathscr{L}_{\tilde X}\tilde G$ of a tensor field $\tilde G$ of type $(0, 2)$ with respect to a vector field $\tilde X$ is defined by

$$(\mathscr{L}_{\tilde X}\tilde G)(\tilde Y, \tilde Z) = \tilde X(\tilde G(\tilde Y, \tilde Z)) - \tilde G([\tilde X, \tilde Y], \tilde Z) - \tilde G(\tilde Y, [\tilde X, \tilde Z]),$$

$\tilde Y$ and $\tilde Z$ being arbitrary vector fields.

Proposition 6.2. *For any $A, B, C \in \mathfrak{gl}(n)$, $\alpha, \beta \in S^2(n)$ and X, Y vector fields on M, we have*

(1) $(\mathcal{L}_{\lambda A} G^D)(X^H, \lambda B) = (\mathcal{L}_{\lambda A} G^D)(\lambda B, X^H) = 0$,

(2) $(\mathcal{L}_{\lambda A} G^D)(X^H, \lambda\beta) = (\mathcal{L}_{\lambda A} G^D)(\lambda\beta, X^H) = 0$,

(3) $(\mathcal{L}_{\lambda A} G^D)(X^H, Y^H) = 0$,

(4) $(\mathcal{L}_{\lambda A} G^D)(\lambda\beta, \lambda C) = G^\circ(B(A+A^t), C)$,

(5) $(\mathcal{L}_{\lambda A} G^D)(\lambda B, \lambda\beta) = (\mathcal{L}_{\lambda A} G^D)(\lambda\beta, \lambda B) = 0$,

(6) $(\mathcal{L}_{\lambda A} G^D)(\lambda\alpha, \lambda\beta) = G^\circ(\alpha\circ(I, A+A^t), \beta) + G^\circ(\alpha, \beta^\circ(I, A+A^t))$,

where A^t denotes the transpose of A.

Proof. (1), (2), (3) and (5) follow directly from (3.2), (3.9) and (6.4). (4) and (6) follow by a direct computation using (6.2) and (6.3).

Corollary 6.3. *Let G be a Riemannian metric (resp., an almost symplectic form) on M. Then the fundamental vector field λA on $\mathcal{F}^2 M$ is a Killing vector field (resp., an infinitesimal automorphism) of $(\mathcal{F}^2 M, G^D)$ if and only if $A + A^t = 0$, that is, if and only if A is skewsymmetric.*

Proposition 6.4. *For any $A, B \in \mathfrak{gl}(n)$, $\alpha, \beta, \gamma \in S^2(n)$ and X, Y vector fields on M, we have*

(1) $(\mathcal{L}_{\lambda\alpha} G^D)(X^H, \lambda B) = (\mathcal{L}_{\lambda\alpha} G^D)(\lambda B, X^H) = 0$,

(2) $(\mathcal{L}_{\lambda\alpha} G^D)(X^H, \lambda\beta) = (\mathcal{L}_{\lambda\alpha} G^D)(\lambda\beta, X^H) = 0$,

(3) $(\mathcal{L}_{\lambda\alpha} G^D)(X^H, Y^H) = 0$,

(4) $(\mathcal{L}_{\lambda\alpha} G^D)(\lambda A, \lambda B) = 0$,

(5) $(\mathcal{L}_{\lambda\alpha} G^D)(\lambda A, \lambda\beta) = (\mathcal{L}_{\lambda\alpha} G^D)(\lambda\beta, \lambda A) = 0$,

(6) $(\mathcal{L}_{\lambda\alpha} G^D)(\lambda\beta, \lambda\gamma) = 0$.

Proof. It follows by a direct computation in a way similar to that of the proof of Proposition 6.2.

Corollary 6.5. *Let G be a Riemannian metric (resp., an almost symplectic form) on M. Then the fundamental vector field $\lambda\alpha$ on $\mathcal{F}^2 M$ is always a Killing vector field (resp., an infinitesimal automorphism) of $(\mathcal{F}^2 M, G^D)$.*

From Corollary 6.3 and Corollary 6.5, we have

Corollary 6.6. *Let G be a Riemannian metric (resp., an almost symplectic form) on M. Then the fundamental vector field $\lambda(A, \alpha)$ on $\mathcal{F}^2 M$ is a Killing vector* *field (resp., an infinitesimal automorphism) of* $(\mathscr{F}^2M, G^D)$ *if and only if* $A+A^t=0$, *that is, A is skewsymmetric.*

Proposition 6.7. *For any* $A, B \in \mathrm{gl}(n)$, $\alpha, \beta \in S^2(n)$ *and vector fields X, Y, Z on M, we have*

(1) $(\mathscr{L}_{X^H} G^D)(Y^H, Z^H) = ((\mathscr{L}_X G)(Y, Z))^V$,

(2) $(\mathscr{L}_{X^H} G^D)(Y^H, \lambda A) = G^D(\lambda R(X, Y)^\circ, \lambda A)$,

$(\mathscr{L}_{X^H} G^D)(\lambda A, Y^H) = G^D(\lambda A, \lambda R(X, Y)^\circ)$,

(3) $(\mathscr{L}_{X^H} G^D)(Y^H, \lambda\alpha) = 2G^D(\lambda\Omega_1(X^H, Y^H), \lambda\alpha)$,

$(\mathscr{L}_{X^H} G^D)(\lambda\alpha, Y^H) = 2G^D(\lambda\alpha, \lambda\Omega_1(X^H, Y^H))$,

(4) $(\mathscr{L}_{X^H} G^D)(\lambda A, \lambda B) = (\nabla_X G)^\circ(A, B)$,

(5) $(\mathscr{L}_{X^H} G^D)(\lambda A, \lambda\alpha) = (\mathscr{L}_{X^H} G^D)(\lambda\alpha, \lambda A) = 0$,

(6) $(\mathscr{L}_{X^H} G^D)(\lambda\alpha, \lambda\beta) = (\nabla_X G)^\circ(\alpha, \beta)$.

Proof. The proof follows by a straightforward computation from (3.2), (3.9), (6.2), (6.3) and (6.4).

Corollary 6.8. *Let X be a vector field and G a tensor field of type* (0, 2) *on M. Then the condition* $\mathscr{L}_{X^H} G^D=0$ *is equivalent to the conditions*

$$\mathscr{L}_X G = 0, \quad R(X, \cdot) = 0, \quad \nabla_X G = 0, \quad i_{X^H}\Omega_1 = 0,$$

where $R(X, \ )$ *denotes the tensor field of type* (1, 2) *on M given by* $R(X, \ )(Y, Z)= =R(X, Y)Z$, *for any vector fields Y, Z on M.*

Let us now suppose that the induced connection ∇ on M from a connection Γ of order 2 is the Riemannian connection of a Riemannian metric G on M. Then, we have

Theorem 6.9. *If the horizontal lift* X^H *to* $\mathscr{F}^2M$ *of a vector field X on M is a Killing vector field in* $(\mathscr{F}^2M, G^D)$, *then X is a Killing vector field in* (M, G). *Conversely, suppose that Γ is partially flat and X is a Killing vector field with vanishing second covariant derivative in* (M, G); *then* X^H *is a Killing vector field in* $(\mathscr{F}^2M, G^D)$.

Proof. We only need to prove the converse. Indeed, if X is a Killing vector field in (M, G), then $\mathscr{L}_X G=0$ and, hence $\mathscr{L}_X\nabla=0$. Therefore, $R(X, Y)=-\nabla_Y(\nabla X)$. Moreover, since $(\nabla_Y(\nabla X))(Z)=(\nabla^2 X)(Z, Y)$, we have $R(X, Y)Z= =-(\nabla^2 X)(Z, Y)$.

To end this section, we shall consider an almost symplectic form G on M. The following set of formulas is easily obtained by a straightforward computation:

$$
\begin{aligned}
&dG^D(X^H, Y^H, Z^H) = \{dG(X, Y, Z)\}^V,\\
&dG^D(X^H, Y^H, \lambda C) = (1/3)G^D(\lambda(R(X, Y)^\circ), \lambda C),\\
&dG^D(X^H, Y^H, \lambda\gamma) = (2/3)G^D(\lambda(\Omega_1(X^H, Y^H)), \lambda\gamma),\\
&dG^D(X^H, \lambda B, \lambda C) = (1/3)\{(\nabla_X G)^\circ(B, C)\},\\
&dG^D(X^H, \lambda B, \lambda\gamma) = 0,\\
(6.5)\qquad &dG^D(X^H, \lambda\beta, \lambda\gamma) = (1/3)\{(\nabla_X G)^\circ(\beta, \gamma)\},\\
&dG^D(\lambda A, \lambda B, \lambda C) = 0,\\
&dG^D(\lambda A, \lambda B, \lambda\gamma) = 0,\\
&dG^D(\lambda A, \lambda\beta, \lambda\gamma) = (1/3)\{G^\circ(\beta\circ(I, A)+\beta\circ(A, I), \gamma)+G^\circ(\beta, \gamma\circ(I, A)+\gamma\circ(A, I))\},\\
&dG^D(\lambda\alpha, \lambda\beta, \lambda\gamma) = 0
\end{aligned}
$$

for any vector fields X, Y, Z on M, any $A, B \in \mathrm{gl}(n)$ and any $\alpha, \beta, \gamma \in S^2(n)$.

Then, we have

Proposition 6.10. *The almost symplectic form G is never closed; consequently, the almost symplectic manifold $(\mathscr{F}^2 M, G^D)$ is never sympletic.*

Proof. In fact, if we take $A=I$ in (6.5), we obtain

$$dG^D(\lambda I, \lambda\beta, \lambda\gamma) = (4/3)G^\circ(\beta, \gamma) \quad \text{for any} \quad \beta, \gamma \in S^2(n).$$

7. The frame bundle of order 2 of an almost Hermitian manifold

Let M be an m-dimensional manifold, J a tensor field on M of type (1, 1) such that $J^2=-I$ and G a Riemannian metric on M such that $G(JX, JY)=G(X, Y)$ for any vector fields X, Y on M; then, (M, J, G) is said to be an almost Hermitian manifold.

Let Γ be a connection of order 2 on M. Since $(J^D)^2=-I$ from Proposition 5.3 and G^D is a Riemannian metric on $\mathscr{F}^2 M$ from Proposition 6.1, we have

Proposition 7.1. *$(\mathscr{F}^2 M, J^D, G^D)$ is an almost Hermitian manifold.*

Proof. It suffices to check the identity

$$G^D(J^D\tilde{X}, J^D\tilde{Y}) = G^D(\tilde{X}, \tilde{Y})$$

in the following three particular cases:

(1) $\tilde{X}=X^H$, $\tilde{Y}=Y^H$, X, Y being arbitrary vector fields on M. The identity follows taking into account Proposition 5.1 and (6.4).

(2) $\tilde{X}=X^H$ and $\tilde{Y}=\lambda(A, \alpha)$ for any vector field X on M and any $A\in \mathrm{gl}(n)$, $\alpha\in S^2(n)$. In this case, both members of the identity vanish from Proposition 5.1 and (6.4).

(3) $\tilde{X}=\lambda(A, \alpha)$, $\tilde{Y}=\lambda(B, \beta)$ for any $A, B\in \mathrm{gl}(n)$, $\alpha, \beta\in S^2(n)$. The result follows by a straightforward computation taking into account Proposition 5.1, (6.4) and the Hermitian character of G.

Let Φ be the Kaehler form of (M, J, G), that is, Φ is the 2-form on M given by

$$\Phi(X, Y) = G(X, JY),$$

for any vector fields X, Y on M. Then, we have

Proposition 7.2. *The Kaehler form of the almost Hermitian manifold* $(\mathcal{F}^2M, J^D, G^D)$ *is the diagonal lift* Φ^D *of the Kaehler form* Φ *of* (M, J, G).

Proof. The proof is similar to that of Proposition 7.1 and is left to the reader.

Now, let us recall that an almost Hermitian manifold (M, J, G) is said to be 1) Hermitian, if $N_J=0$; 2) almost Kaehler, if $d\Phi=0$; 3) Kaehler, if $N_J=0$ and $d\Phi=0$. Moreover, it is well known that the Kaehler form Φ of (M, J, G) is almost symplectic. Thus, we deduce

Theorem 7.3. *Let* (M, J, G) *be an almost Hermitian manifold. Then*

(1) $(\mathcal{F}^2M, J^D, G^D)$ *is never an almost Kaehler manifold.*

(2) *Moreover, if* (M, J, G) *is a Kaehler manifold and its Riemannian connection is the induced linear connection on* M *from* Γ, *then* $(\mathcal{F}^2M, J^D, G^D)$ *is an Hermitian manifold if* Γ *has zero curvature.*

Proof. (1) is a direct consequence of Proposition 6.10, and (2) follows easily from Proposition 5.5.

References

[1] L. A. Cordero and M. de Leon, Lifts of tensor fields to the frame bundle, *Rend. Circ. Mat. Palermo*, 32 (1983), 236—271.

[2] J. Gancarzewicz, Geodesics of order 2, *Zeszyty Nauk. Uniw. Jagielloń. Prace Mat.*, 19 (1977), 121—136.

[3] J. Gancarzewicz, Complete lifts of tensor fields of type $(1, k)$ to natural bundles, *Zeszyty Nauk. Uniw. Jogielloń. Prace Mat.*, 23 (1982), 51—84.

[4] S. Kobayashi, Frame bundles of higher order contact, in: *Differential Geometry*, Proc. Symp. Pure Math., Vol. 3, Amer. Math. Soc. (1961); pp. 186—193.

[5] S. KOBAYASHI and K. NOMIZU, *Foundations of Differential Geometry*, Vol. 1 and 2, Interscience (New York, 1963—1969).

[6] S. KOBAYASHI, *Transformation groups in Differential Geometry*, Springer (Berlin, 1972).

[7] M. DE LEON and M. SALGADO, G-structures on the frame bundle of second order, *Riv. Mat. Univ. Parma*, (4) **11** (1985).

DEPARTAMENTO DE GEOMETRÍA Y TOPOLOGÍA
FACULTAD DE MATEMÁTICAS
UNIVERSIDAD DE SANTIAGO DE COMPOSTELA
SANTIAGO DE COMPOSTELA, SPAIN

Acta Sci. Math., 50 (1986), 87—92

On norms of projections

A. P. BOSZNAY and B. M. GARAY

Let $(X, \|\cdot\|)$ be a normed space. A continuous linear mapping $P: X\to X$ is said to be a projection if $P^2=P$. As usual, the range and the null space of P is denoted by $\mathscr{R}(P)$ and $\mathscr{N}(P)$, respectively. Further, the norm of P is defined as $\|P\|=\sup\{\|Px\| \,|\, \|x\|\leqq 1\}$. Clearly $\|P\|\geqq 1$ excepting $P=0$ and $\|I\|=1$. (Here 0 and I denotes the zero and the identity operator on X, respectively.)

Let Y be a one-dimensional subspace of X. It follows immediately from Hahn—Banach theorem [3] that there exists a projection $P: X\to X$ for which $\mathscr{R}(P)=Y$ and $\|P\|=1$.

The aim of this paper is to investigate the question of the existence of normed spaces for which $P\neq I$ and $\dim\mathscr{R}(P)>1$ imply $\|P\|>1$.

By a density theorem, we solve the problem in finite dimensions. The infinite dimensional case seems to be entirely open.

From now on, let X be an n-dimensional real vector space. Assume that $n\geqq 3$. Let $\langle\cdot,\cdot\rangle: X\times X\to\mathbf{R}$ be a fixed scalar product. Let $e_1, \ldots, e_n$ be a fixed orthonormal system with respect to the scalar product $\langle\cdot,\cdot\rangle$. Every $x\in X$ has a unique representation of the form $x=\sum_{i=1}^{n}\alpha_i e_i$; $\alpha_i\in\mathbf{R}$. Thus the basis $e_1, \ldots, e_n$ determines a one to one correspondence between vectors of X and n-tuples (column vectors) in $\mathbf{R}^n$. Given $x_1, \ldots, x_n\in X$, now it is possible to define the determinant function $\det(x_1, \ldots, x_n)$, as a function of column vectors in $\mathbf{R}^n$.

The set of norms defined on X will be denoted by $N(X)$. $N(X)$ can be made into a metric space in a very natural way. The distance between two norms $\|\cdot\|_1, \|\cdot\|_2: X\to\mathbf{R}^+$ can be defined as $d(\|\cdot\|_1, \|\cdot\|_2)=\sup\{|\|x\|_1-\|x\|_2| \,\big|\, \langle x, x\rangle=1\}$. As any two scalar products (moreover, any two norms) induce the same topology on X, the topology on $N(X)$ induced by d does not depend on the particular choice of the scalar product $\langle\cdot,\cdot\rangle$. Therefore, we are justified in speaking about openness and density of subsets of X as well as of $N(X)$ without referring to any particular scalar product.

Received January 10, 1984.

Now we are in a position to formulate our main result.

Theorem. *Let X be an n-dimensional real vector space. Assume that* $n\geqq 3$. *Define*

$$N_1(X) =$$

$$= \{\|\cdot\|\in N(X) \mid \textit{for any projection } P\colon X\to X;\ P\neq I \textit{ and } \dim \mathscr{R}(P)>1 \textit{ imply } \|P\|>1\}.$$

Then $N_1(X)$ is open and dense in $N(X)$.

Proof of the openness of $N_1(X)$. Pick a norm $\|\cdot\|_1$ in $N_1(X)$. There exists a constant $K>0$ such that $(\langle x, x\rangle)^{1/2}\leqq K\|x\|_1$. If $d(\|\cdot\|_1, \|\cdot\|_2)\leqq\eta$, then $|\|x\|_1-\|x\|_2|\leqq\eta(\langle x, x\rangle)^{1/2}$, and consequently, $(1-\eta K)\|x\|_1\leqq\|x\|_2\leqq(1+\eta K)\|x\|_1$.

It follows easily from a compactness argument that $\inf\{\|P\|_1 \mid P\colon X\to X$ is a projection, $P\neq I$, $\dim\mathscr{R}(P)>1\}>1$, i.e. for any projection $P\colon X\to X$ there holds $\|P\|_1\geqq 1+\alpha$ for some fixed $\alpha>0$ provided that $P\neq I$, $\dim\mathscr{R}(P)>1$. Therefore, $\|Px\|_2\geqq(1-\eta K)\|Px\|_1\geqq(1-\eta K)(1+\alpha)\|x\|_1\geqq(1-\eta K)(1+\alpha)(1+\eta K)^{-1}\|x\|_2$. Consequently, η being sufficiently small implies $\|P\|_2>1$, $\|\cdot\|_2\in N_1(X)$.

The proof of the density of $N_1(X)$ requires more difficult considerations. If $S\subset X$, the set of all linear combinations of elements of S, i.e. the subspace spanned by S is denoted by $\operatorname{Span}(S)$. The orthogonal complement of $\operatorname{Span}(S)$ is denoted by $\operatorname{Span}^\perp(S)$. Let us recall that $\dim\operatorname{Span}(S)+\dim\operatorname{Span}^\perp(S)=n$.

Definition. Let N be a fixed positive integer. For sake of brevity, we call a set $\{x_1, \dots, x_N\}\subset X$ to be independent if for any $Y\subset\{x_1, \dots, x_N\}$ and for any partition $Y_1=\{{}^1x_1, \dots, {}^1x_{n_1}\}, \dots, Y_k=\{{}^kx_1, \dots, {}^kx_{n_k}\}$ of Y ($k\geqq 1$, $Y_i\cap Y_j=\emptyset$ if $i\neq j$; $i, j=1, \dots, k$; $n_j\geqq 0$; $\bigcup_{j=1}^{k} Y_j=Y$) satisfying $n_j\leqq n$, $j=1, \dots, k$ there holds

$$\dim\Big(\bigcap_{j=1}^{k}\operatorname{Span}(Y_j)\Big) = \max\Big\{0, n-\sum_{j=1}^{k}(n-n_j)\Big\}. \tag{1}$$

Further, we say that a vector $\bar{x}=(x_1, \dots, x_N)\in X^N$ is of type $\mathscr{I}$ if the set of its coordinate vectors $\{x_1, \dots, x_N\}$ is independent.

Remark 1. In case of $N=n$, $k=1$ one arrives at the usual definition of linear independence.

Remark 2. On the account of

$$\Big(\bigcap_{j=1}^{k}\operatorname{Span}(Y_j)\Big)^\perp = \operatorname{Span}(\{\operatorname{Span}^\perp(Y_1), \dots, \operatorname{Span}^\perp(Y_k)\}),$$

(1) can be reformulated as

$$\dim(\operatorname{Span}(\{\operatorname{Span}^\perp(Y_1), \dots, \operatorname{Span}^\perp(Y_k)\})) = \min\Big\{n, \sum_{j=1}^{k}(n-n_j)\Big\}. \tag{1'}$$

Lemma 1. *The set of vectors of type $\mathcal{S}$ is dense in X^N.*

Proof. Pick a vector $\bar{x}=(x_1, \dots, x_N)\in X^N$. Consider a partition $Y_1=\{{}^1x_1, \dots, {}^1x_{n_1}\}, \dots, Y_k=\{{}^kx_1, \dots, {}^kx_{n_k}\}$ of a subset $Y\subset\{x_1, \dots, x_N\}$ satisfying $n_j\leqq n, j=1, \dots, k$.

For $j=1, \dots, k$; $l=1, \dots, n-n_j$ define

$${}^jw_l = \det({}^jx_1, \dots, {}^jx_{n_j}, e_1, \dots, e_{l-1}, e_{l+1}, \dots, e_{n-n_j}, \operatorname{col}(e_1, \dots, e_n)).$$

The vectors ${}^jw_1, \dots, {}^jw_{n-n_j}$ form a basis for $\operatorname{Span}^\perp(Y_j)$ provided that

$$(2) \qquad \det({}^jx_1, \dots, {}^jx_{n_j}, {}^jw_1, \dots, {}^jw_{n-n_j}) \neq 0.$$

Suppose that any $p\leqq n$ vectors in $\bigcup_{j=1}^{k}\{{}^jw_1, \dots, {}^jw_{n-n_j}\}$, say $z_1, \dots, z_p$ satisfy

$$(3) \qquad \det(z_1, \dots, z_p, e_1, \dots, e_{n-p}) \neq 0.$$

It is obvious that (2) and (3) imply (1'). Thus the set of vectors of type $\mathcal{S}$ contains the complement of a real algebraic variety, consequently [4], it is dense in X^N.

Suppose now $x_1,\dots, x_n$ is a basis of X. The $(n-1)$-simplex $\sigma[x_1, \dots, x_n]$ with vertices $x_1, \dots, x_n$ and its interior $\operatorname{int}(\sigma[x_1, \dots, x_n])$ is defined as the set of all $x\in X$ of the form

$$x = \sum_{i=1}^{n}\alpha_i x_i \quad \text{where} \quad \alpha_i \geqq 0; \quad i = 1, \dots, n; \quad \sum_{i=1}^{n}\alpha_i = 1$$

and

$$x = \sum_{i=1}^{n}\alpha_i x_i \quad \text{where} \quad \alpha_i > 0; \quad i = 1, \dots, n; \quad \sum_{i=1}^{n}\alpha_i = 1$$

respectively. The vector $v(\sigma[x_1, \dots, x_n])$ defined by

$$\det(x_2-x_1, \dots, x_n-x_1, \operatorname{col}(e_1, \dots, e_n))$$

is a nonzero element in $\operatorname{Span}^\perp(\{x_2-x_1, \dots, x_n-x_1\})$, i.e. it is a normal vector to $\sigma[x_1, \dots, x_n]$.

Lemma 2. *Let $\bar{x}=(x_1, \dots, x_N)\in X^N$ be a vector of type $\mathcal{S}$. Given $\varepsilon>0$, then there exists an N-tuple of real numbers $(\varepsilon_1, \dots, \varepsilon_N)\in\mathbf{R}^N$ satisfying $0<\varepsilon_l<\varepsilon$; $l=1, \dots, N$ such that for any $Z\subset\{x_1, \dots, x_N, -x_1, \dots, -x_N\}$ and for any partition $Z_1=\{{}^1x_1, \dots, {}^1x_n\}, \dots, Z_n=\{{}^nx_1, \dots, {}^nx_n\}$ of Z ($Z_i\cap Z_j=\emptyset$ if $i\neq j$; $i,j=1, \dots, n$; $\bigcup_{j=1}^{n} Z_j=Z$) satisfying*

$${}^ix_j \neq \pm{}^px_r \quad \textit{for} \quad |i-p|+|j-r| \neq 0$$

there holds

(4) $$\operatorname{Span}\big(\{v(\sigma[(1+{}^1\varepsilon_1)^1x_1, \ldots, (1+{}^1\varepsilon_n)^1x_n]), \ldots \\ \ldots, v(\sigma[(1+{}^n\varepsilon_1)^n x_1, \ldots, (1+{}^n\varepsilon_n)^n x_n])\}\big) = X.$$

Proof. (4) is equivalent to

(4′) $$\det\big(v(\sigma[(1+{}^1\varepsilon_1)^1x_1, \ldots, (1+{}^1\varepsilon_n)^1x_n]), \ldots \\ \ldots, v(\sigma[(1+{}^n\varepsilon_1)^n x_1, \ldots, (1+{}^n\varepsilon_n)^n x_n])\big) \neq 0.$$

As in the proof of Lemma 1, the desired result follows from [4].

A bounded set $K\subset X$ is said to be a centrally symmetric convex polyhedron if there exist nonzero linear functionals $f_s\colon X\to\mathbf{R}$, $s=1, \ldots, t$ such that $K=\bigcap_{s=1}^{t}\{x\in X\,|\,|f_s(x)|\leq 1\}$. The bounding hyperplanes of K are defined as $\{x\in X\,|\,f_s(x)=1\}$, $\{x\in X\,|\,f_s(x)=-1\}$, $s=1, \ldots, t$. Assume that

(5) for any bounding hyperplane H, the intersection $H\cap K$ is an $(n-1)$-simplex. (Such simplices are called the facets of K.)

For sake of brevity, we say that the facets $\sigma[\tilde{x}_1, \ldots, \tilde{x}_n]$ and $\sigma[\tilde{\tilde{x}}_1, \ldots, \tilde{\tilde{x}}_n]$ are non-neighbouring if there holds

$$\{\tilde{x}_1, \ldots, \tilde{x}_n, -\tilde{x}_1, \ldots, -\tilde{x}_n\}\cap\{\tilde{\tilde{x}}_1, \ldots, \tilde{\tilde{x}}_n, -\tilde{\tilde{x}}_1, \ldots, -\tilde{\tilde{x}}_n\} = \emptyset.$$

If $M\subset X$ is a symmetric (i.e. symmetric with respect to the origin) convex set with the origin in its interior, then its Minkowsky functional $\Phi_M\colon X\to\mathbf{R}^+$ defined by $\Phi_M(x)=\inf\{\alpha>0\,|\,x\in\alpha M\}$ is a norm. Conversely, if we are given a norm in X, then the unit ball it defines is a symmetric convex set with the origin in its interior, and it is the corresponding Minkowsky functional.

Proof of the density of $N_1(X)$.

Step 1. Pick a norm $\|\cdot\|$ in $N(X)$. Given $\varepsilon>0$, then there exists a norm $\|\cdot\|_K$ in $N(X)$ with the following properties:

(6) $d(\|\cdot\|, \|\cdot\|_K)<\varepsilon$;

(7) $\|\cdot\|_K=\Phi_K$, the Minkowsky functional of a centrally symmetric convex polyhedron K satisfying (5);

(8) denoting the vertices of K by $x_1, \ldots, x_N$, $-x_1, \ldots, -x_N$, the vector $\bar{x}=(x_1, \ldots, x_N)\in X^N$ is of type $\mathscr{S}$;

(9) if $\sigma_1, \ldots, \sigma_n$ are non-neighbouring facets of K, there holds

$$\operatorname{Span}\big(\{v(\sigma_1), \ldots, v(\sigma_n)\}\big) = X;$$

and

(10) for any two-dimensional (linear) subspace $W\subset X$, the number of pairwise non-neighbouring facets of K intersecting W at a segment, is at least $n(n-1)$.

The existence of $\|\cdot\|_K$ satisfying (6)—(9) follows from the lemmas. (10) is automatically satisfied if

$$\max\{\|\tilde{x}-\tilde{\tilde{x}}\| \mid \text{there exists a facet } \sigma \text{ of } K \text{ such that } \tilde{x}, \tilde{\tilde{x}}\in\sigma\}$$

is sufficiently small.

Step 2. We show that $\|\cdot\|_K\in N_1(X)$. Let us observe first that (8) implies the following property of K:

(11) if the facets $\sigma_1, \ldots, \sigma_{n-1}$ are pairwise non-neighbouring and a two-dimensional (linear) subspace $W\subset X$ intersects each of them at a segment, then, for some $k^*\in\{1, \ldots, n-1\}$, W intersects $\operatorname{int}(\sigma_{k^*})$.

To the contrary, let us suppose that there exists an $l(k)\in\{1, \ldots, n\}$ such that

$$W \subset Y_k = \operatorname{Span}(\{\overset{k}{\smile}x_1, \ldots, \overset{k}{\smile}x_{l(k)-1}, \overset{k}{\smile}x_{l(k)+1}, \ldots, \overset{k}{\smile}x_n\}),$$

for each $k=1, \ldots, n-1$. Since $\dim(\operatorname{Span}(Y_k))=n-1$, (1) yields

$$2 = \dim W \leqq \dim\big(\bigcap_{k=1}^{n-1} \operatorname{Span}(Y_k)\big) = \max\big\{0, n-\sum_{k=1}^{n-1} 1\big\} = 1,$$

a contradiction.

Step 3. Let us suppose now that $P\colon X\to X$ is a projection for which $\|P\|_K=1$, $\dim \mathcal{R}(P)>1$. We have to show that $P=I$.

Consider a two-dimensional (linear) subspace $W\subset\mathcal{R}(P)$. Assume that for a facet σ of K there holds $W\cap\operatorname{int}(\sigma)\neq\emptyset$. We claim that $v(\sigma)\in\mathcal{N}^\perp(P)$. Pick a $z\in W\cap\operatorname{int}(\sigma)$. It is sufficient to show that $x\in\mathcal{N}(P)$ implies $x\in\operatorname{Span}^\perp(v(\sigma))$. In fact, we have $\|z\|_K=\|P(z+\lambda x)\|_K\leqq\|z+\lambda x\|_K$ for arbitrary $\lambda\in\mathbf{R}$. On the other hand, $z\in\operatorname{int}(\sigma)$ implies $(z+\lambda x)\in\sigma$ for $|\lambda|$ sufficiently small. Consequently, $x=((z+\lambda x)-z)/\lambda\in\operatorname{Span}^\perp(v(\sigma))$.

By the same reasoning, (10) and (11) imply the existence of pairwise non-neighbouring facets $\sigma_1, \ldots, \sigma_n$ of K such that $v(\sigma_1), \ldots, v(\sigma_n)\in\mathcal{N}^\perp(P)$. Applying (9) we obtain $X\subset\mathcal{N}^\perp(P)$, which, in turn, implies that $P=I$.

For applications of the Theorem, see [1], [2].

Remark 3. The Theorem remains valid if X is allowed to be a complex finite dimensional vector space.

The following problems arise naturally:

Problem 1. What is the minimum number of vertices of centrally symmetric convex polyhedra satisfying $\|\cdot\|_K\in N_1(X)$? (In the three-dimensional real case it is not hard to construct a centrally symmetric convex polyhedron K with twelve vertices for which $\|\cdot\|_K\in N_1(X)$. On the other hand, it seems plausible that there are no such polyhedra with ten vertices. Nevertheless, we are not able to prove it.)

Problem 2. Give upper and lower bounds for

$$\sup\{\inf\{\|P\| \mid P\colon X \to X \text{ is a projection satisfying } \dim \mathcal{R}(P) > 1,\ P \neq I\} \mid \|\cdot\| \in N_1(X)\}.$$

Problem 3. The infinite dimensional case.

References

[1] A. P. Bosznay and B. M. Garay, A geometric problem concerning discs, *Acta Sci. Math.*, to appear.

[2] B. M. Garay and V. Kertész, Lozinsky's functional and the transformation of variables, *Z. Anal. Anw.*, **2** (1984), 87—95.

[3] W. Rudin, *Functional Analysis*, McGraw-Hill (New York, 1973).

[4] H. Whitney, Elementary structure of real algebraic varieties, *Ann. Math.*, **66** (1957), 545—556.

DEPARTMENT OF MATHEMATICS
BUDAPEST UNIVERSITY OF TECHNOLOGY
MŰEGYETEM RKP. 3—9
1521 BUDAPEST, HUNGARY

Acta Sci. Math., 50 (1986), 93—98

On relations of coefficient conditions

V. TOTIK and I. VINCZE

In the theory of orthogonal series several different kinds of coefficient condition are being used and among them the three most frequently investigated conditions are

(i) $$\sum_{n=1}^{\infty} c_n^2 \varrho_n < \infty,$$

(ii) $$\sum_{k=1}^{\infty} \lambda_k \Big(\sum_{n=\nu_k+1}^{\nu_{k+1}} c_n^2 \Big)^{\varepsilon/2} < \infty$$

and

(iii) $$\sum_{k=1}^{\infty} \varkappa_k \Big(\sum_{n=k}^{\infty} c_n^2 \Big)^{\varepsilon/2} < \infty.$$

Here $\varepsilon > 0$, $\{\varrho_n\}$, $\{\lambda_n\}$ and $\{\varkappa_n\}$ are certain monotone sequences of real numbers and $\{c_n\}$ is a real coefficient sequence. For different results incorporating (i), (ii) or (iii) we refer to [2], a paper devoted to the systematic study of the connections between (i), (ii) or (iii). In it L. Leindler gave sufficient conditions for the equivalence of (ii) and (iii) (for any sequence $\{c_n\}$) and investigated the relation between (i) and (ii). Our aim in this paper is to give necessary and sufficient conditions for the equivalences above in a somewhat more general setting.

Let us consider the conditions

(1) $$\sum_{n=1}^{\infty} \varrho_n |c_n|^q < \infty$$

(2) $$\sum_{k=1}^{\infty} \lambda_k \Big(\sum_{n=\nu_k+1}^{\nu_{k+1}} |c_n|^q \Big)^{p/q} < \infty$$

and

(3) $$\sum_{k=1}^{\infty} \varkappa_k \Big(\sum_{n=k}^{\infty} |c_n|^q \Big)^{p/q} < \infty,$$

Received November 21, 1983.

where p and q are positive numbers, $\{\varrho_n\}$ and $\{\varkappa_k\}$ positive sequences, $\{\lambda_k\}$ is a positive monotone sequence and $\{\nu_k\}$ is a subsequence of the natural numbers.

We have the following theorems:

Theorem 1. a) *If $p=q$ then conditions* (2) *and* (3) *are equivalent for every* $\{c_n\}$ *if and only if*

$$(1/A)\sum_{k=1}^{\nu_{m+1}} \varkappa_k \leqq \lambda_m \leqq A\sum_{k=1}^{\nu_m+1} \varkappa_k \quad (m = 1, 2, 3, \ldots) \tag{4}$$

is satisfied with some constant A.

b) *If $p\neq q$ then* (2) *and* (3) *are equivalent if and only if* (4) *and*

$$\lambda_{m+A} \geqq 2\lambda_m \quad (m = 1, 2, \ldots) \tag{5}$$

are satisfied with some natural number A.

Theorem 2. *Conditions* (1) *and* (2) *cannot be equivalent unless $p=q$. If $p=q$ then they are equivalent if and only if there is an A with*

$$(1/A)\lambda_m \leqq \varrho_n \leqq A\lambda_m \quad (m = 1, 2, 3, \ldots;\ \nu_m < n \leqq \nu_{m+1}).$$

Theorem 3. a) *If $p\neq q$ then* (1) *and* (3) *are equivalent if and only if the three sequences*

$$\{\varrho_n\},\quad \{1/\varrho_n\},\quad \Big\{\sum_{k=1}^{n} \varkappa_k\Big\}$$

are bounded.

b) *If $p=q$ then* (1) *and* (3) *are equivalent if and only if*

$$(1/A)\varrho_m \leqq \sum_{k=1}^{m} \varkappa_k \leqq A\varrho_m \quad (m = 1, 2, 3, \ldots)$$

is satisfied with a constant A.

Our theorems have several consequences, the most remarkable one is that, since e.g. (4) and (5) are independent of p and q $(p\neq q)$, the equivalence of (2) and (3) for a pair p, q $(p\neq q)$ implies their equivalence for any other pair p', q'. Another direct corollary of Theorem 1 is [2, Theorem 2.1].

The sufficiency of our conditions can be verified by more or less direct considerations using some well-known inequalities (such as Jensen's inequality) from the theory of sequences. Our necessity proofs, however, have a very general character — a certain boundedness principle is applied in them. Since the proofs run on similar lines we shall give a detailed proof only for Theorem 1. However, we emphasize that the method can be applied to Theorems 2 and 3 and to other equivalence problems of this kind.

Proof of Theorem 1. We separately prove the necessity and sufficiency of our conditions.

I. Necessity. Let X_1 and X_2 be the set of the sequences $\{c_n\}$ for which (2) and (3) are satisfied, respectively. Then with the usual operations X_1 and X_2 are linear spaces. We introduce on them length functions $\| \|_i$: $X_i \to R_+$ $(i=1, 2)$ as follows: for $c=\{c_n\}$ let

$$|c|_1 = \sum_{k=1}^{\infty} \lambda_k \Big(\sum_{n=\nu_k+1}^{\nu_{k+1}} |c_n|^q \Big)^{p/q},$$

$$\|c\|_1 = \begin{cases} (|c|_1)^{1/p} & \text{if } p, q \geqq 1, \\ |c|_1 & \text{if } 0<p\leqq q, \quad 0<p<1, \\ (|c|_1)^{q/p} & \text{if } 0<q\leqq p, \quad 0<q<1, \end{cases}$$

and

$$|c|_2 = \sum_{k=1}^{\infty} \varkappa_k \Big(\sum_{n=k}^{\infty} |c_n|^q \Big)^{p/q},$$

$$\|c\|_2 = \begin{cases} (|c|_2)^{1/p} & \text{if } p, q \geqq 1, \\ |c|_2 & \text{if } 0<p\leqq q, \quad 0<p<1, \\ (|c|_2)^{q/p} & \text{if } 0<q<p, \quad 0<q<1. \end{cases}$$

These length functions induce two metrics:

$$d_1(c_1, c_2) = \|c_1-c_2\|_1, \quad d_2(c_1, c_2) = \|c_1-c_2\|_2$$

and it is easy to see that (X_1, d_1) and (X_2, d_2) are complete metric spaces, the metrics d_1 and d_2 are invariant, i.e. $d_i(c_1, c_2)=d_i(c_1-c_2, 0)$ $(i=1, 2)$, furthermore, the mappings $(\lambda, c)\to\lambda c$ of $R\times X_i \to X_i$ $(i=1, 2)$ are continuous in λ for each c and in c for each λ.

Summarizing and putting into the terminology of function spaces we can say that (X_1, d_1) and (X_2, d_2) are F-spaces (see [1, pp. 50—51]).

Now suppose that (2) implies (3), i.e. $X_1 \subseteqq X_2$. For $c=\{c_n\}\in X_1$ and a natural number m let

$$T_m c = d \quad \text{where} \quad d = \{c_1, c_2, \ldots, c_m, 0, 0, \ldots\}.$$

By the assumption the sequence $\{T_m\}$ of the bounded linear operators T_m: $X_1 \to X_2$ is pointwise bounded, i.e. for every $c\in X_1$ there is a bound K_c such that

$$\|T_m c\|_2 \leqq K_c \quad (m = 1, 2, \ldots).$$

Since the length functions $\|\cdot\|_1$ and $\|\cdot\|_2$ are homogeneous of the same degree, the theorem of Banach and Steinhaus valid for operators between F-spaces (see [1, p. 52]) yields that there exists a uniform bound A such that

$$\|T_m c\|_2 \leqq A \|c\|_1 \quad (c\in X_1, \ m = 1, 2, \ldots)$$

is satisfied; and letting here m tend to infinity we arrive at

$$\|c\|_2 \leqq A\|c\|_1 \quad (c\in X_1).$$

Similarly, it can be proved that if (3) implies (2) then with some constant

$$\|c\|_1 \leqq A\|c\|_2 \quad (c\in X_2).$$

Now we introduce the perhaps somewhat awkward, nonetheless suggestive notations

$$((2)) = \sum_{k=1}^{\infty} \lambda_k \Big(\sum_{n=\nu_k+1}^{\nu_{k+1}} |c_n|^q \Big)^{p/q},$$

$$((3)) = \sum_{k=1}^{\infty} \varkappa_k \Big(\sum_{n=k}^{\infty} |c_n|^q \Big)^{p/q}$$

for the sums involved in (2) and (3).

According to what we have proved above the equivalence of (2) and (3) implies the inequalities $((3))\leqq A((2))$ and $((2))\leqq A((3))$. Applying the first one of these to the sequence

$$c_n = \begin{cases} 1 & \text{if} \quad n = \nu_{m+1} \\ 0 & \text{otherwise} \end{cases}$$

and the second one to

$$c_n = \begin{cases} 1 & \text{if} \quad n = \nu_m + 1 \\ 0 & \text{otherwise} \end{cases}$$

where m is a fixed natural number we get that

$$\sum_{k=1}^{\nu_{m+1}} \varkappa_k \leqq A\lambda_m \quad \text{and} \quad \lambda_m \leqq A \sum_{k=1}^{\nu_m+1} \varkappa_k$$

are satisfied and the necessity of (4) has been verified.

To prove that (5) also holds we remark first that in the case $\lambda_m \searrow$ we have by (4) $\lambda_1 \geqq \lambda_k \geqq (1/A^2)\lambda_1$ and so from our point of view the sequence $\{\lambda_k\}$ is the same as the sequence $\lambda_k^* \equiv 1$ (condition (2) does not change if we replace $\{\lambda_k\}$ by $\{\lambda_k^*\}$). Thus, we may assume $\{\lambda_k\}$ to be nondecreasing (and the proof below shows that $\lambda_m \searrow$ cannot occur at all).

Let m and $s/2$ be two integers. Applying the inequality $((2))\leqq A((3))$ to the sequence

$$c_n = \begin{cases} 1 & \text{if} \quad n = \nu_{m+2}, \quad \nu_{m+3}, \dots, \nu_{m+s} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

we obtain for $p/q<1$ that

$$s\lambda_m \leqq \sum_{k=m}^{m+s-1} \lambda_k = ((2)) \leqq A((3)) \leqq A \sum_{k=1}^{\nu_{m+s}} \varkappa_k s^{p/q} \leqq A^2 s^{p/q} \lambda_{m+s}$$

and so

$$\lambda_{m+s} \geqq s^{1-p/q} \cdot A^{-2} \cdot \lambda_m \geqq 2 \cdot \lambda_m$$

if $s > (2A)^{2(1-p/q)}$, which proves (5).

Similarly, if $p/q > 1$ then we obtain from $((3)) \leqq A((2))$ applied to the sequence (6) that

$$\lambda_m (s/2)^{p/q} \leqq A\Big(\sum_{k=1}^{\nu_m+s/2} \varkappa_k\Big)(s/2) \leqq A((3)) \leqq A^2((2)) = \sum_{k=m}^{m+s-1} \lambda_k \leqq s \cdot \lambda_{m+s}$$

and the proof is over.

II. Sufficiency. Let us assume now (4) and (5) and we shall separately prove that (2) implies (3) and (3) implies (2) for any sequence $\{c_n\}$. Let

$$d_m = \sum_{n=\nu_m+1}^{\nu_{m+1}} |c_n|^q.$$

1) (2) implies (3). Since

$$((3)) \leqq \sum_{m=1}^{\infty} \Big(\sum_{k=\nu_m+1}^{\nu_{m+1}} \varkappa_k\Big)\Big(\sum_{j=m}^{\infty} d_j\Big)^{p/q} = I$$

it is enough to show that (2) implies $I < \infty$.

If $p/q \leqq 1$ then from the concavity of $x^{p/q}$ we get

$$I \leqq \sum_{m=1}^{\infty} \Big(\sum_{k=\nu_m+1}^{\nu_{m+1}} \varkappa_k\Big)\Big(\sum_{j=m}^{\infty} d_j^{p/q}\Big) \leqq \sum_{j=1}^{\infty} \Big(\sum_{k=1}^{\nu_{j+1}} \varkappa_k\Big) d_j^{p/q} \leqq A \sum_{j=1}^{\infty} \lambda_j d_j^{p/q} = A((2)) < \infty.$$

If, however, $p/q > 1$ then we have by (5) the inequalities

$$\sum_{j=1}^{m} \lambda_j^{q/p} \leqq c \cdot \lambda_m^{q/p}; \quad 1 \leqq \sum_{j=m}^{\infty} \lambda_m^{q/p} / \lambda_j^{q/p} \leqq c \quad (m = 1, 2, \ldots)$$

and so Jensen's inequality gives

$$I \leqq A \sum_{m=1}^{\infty} \lambda_m \Big(\sum_{j=m}^{\infty} d_j\Big)^{p/q} = A \sum_{m=1}^{\infty} \Big(\sum_{j=m}^{\infty} (\lambda_j^{q/p} d_j)(\lambda_m/\lambda_j)^{q/p}\Big)^{p/q} \leqq$$

$$\leqq K \sum_{m=1}^{\infty} \Big(\sum_{j=m}^{\infty} \lambda_j d_j^{p/q} (\lambda_m/\lambda_j)^{q/p}\Big) =$$

$$= K \sum_{j=1}^{\infty} d_j^{p/q} \Big(\sum_{m=1}^{j} \lambda_m^{q/p} \lambda_j^{1-q/p}\Big) \leqq K \sum_{j=1}^{\infty} \lambda_j d_j^{p/q} = K((2)) < \infty$$

and this is what we wanted to prove.

2) (3) implies (2). If $p/q \geqq 1$ then we get from the convexity of $x^{p/q}$ and from (4) that

$$((3)) \geqq \sum_{m=1}^{\infty} \Big(\sum_{k=\nu_m+1}^{\nu_{m+1}} \varkappa_k \Big) \Big(\sum_{j=m+1}^{\infty} d_j \Big)^{p/q} \geqq \sum_{m=1}^{\infty} \Big(\sum_{k=\nu_m+1}^{\nu_{m+1}} \varkappa_k \Big) \Big(\sum_{j=m+1}^{\infty} d_j^{p/q} \Big) =$$

$$= \sum_{j=1}^{\infty} d_j^{p/q} \Big(\sum_{m=1}^{j-1} \sum_{k=\nu_m+1}^{\nu_{m+1}} \varkappa_k \Big) \geqq$$

$$\geqq c \sum_{j=1}^{\infty} \lambda_j d_j^{p/q} - \sum_{j=1}^{\infty} \varkappa_{\nu_j+1} d_j^{p/q} \geqq c((2)) - ((3)) \quad (c > 0)$$

and so $(3) \Rightarrow (2)$ follows.

If $p/q < 1$, then (5) is also satisfied and so there is an s with

$$\sum_{k=\nu_m+1}^{\nu_{m+s}} \varkappa_k \geqq 2 \sum_{k=1}^{\nu_m} \varkappa_k \quad (m = 1, 2, \ldots).$$

Thus, for

$$\gamma_m = \sum_{k=\nu_{ms}+1}^{\nu_{(m+1)s}} \varkappa_k$$

we have

$$\gamma_{m+1} \geqq 2\gamma_m \quad (m = 1, 2, \ldots). \tag{7}$$

Assuming (3), (2) will surely hold if

$$\sum_{m=2}^{\infty} \Big(\sum_{j=1}^{m-1} \gamma_j \Big) \Big(\sum_{l=ms+1}^{(m+1)s} d_l^{p/q} \Big)$$

is finite (take also into account (4)) and by (7) this amounts to the finiteness of

$$\sum_{m=2}^{\infty} \gamma_{m-1} \Big(\sum_{l=ms+1}^{(m+1)s} d_l^{p/q} \Big)$$

which clearly follows from (3).

We have completed our proof.

References

[1] L. Dunford and J. T. Schwartz, *Linear Operators. I,* Interscience Publishers, Inc. (New York, 1958).

[2] L. Leindler, On relations of coefficient conditions, *Acta Math. Acad. Sci. Hungar.*, 39 (1982), 409—420.

BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

Acta Sci. Math., 50 (1986), 99—104

On absolute summability of orthogonal series

L. LEINDLER and K. TANDORI

1. In [7] the second of the authors investigated some questions of general absolute summability of orthogonal series. The aim of the present note is to continue these investigations.

In [6] the following theorem was proved:

Theorem A. *If*

$$\sum_{n=0}^{\infty}\Big\{\sum_{k=2^n+1}^{2^{n+1}} a_k^2\Big\}^{1/2} < \infty \tag{1}$$

holds true, then for any orthonormal system $\{\varphi_n(x)\}$ *on* (a, b) *the orthogonal series*

$$\sum_{k=0}^{\infty} a_k \varphi_k(x) \tag{2}$$

is absolutely Cesàro summable (or briefly $|C, 1|$*-summable) almost everywhere in* (a, b). *If* (1) *does not hold then there exists an orthonormal system* $\{\varphi_n(x)\}$ *such that series* (2) *is not* $|C, 1|$*-summable almost everywhere in* (a, b).

Moreover P. BILLARD [1] proved the following result.

Theorem B. *If the coefficient-sequence* $\{a_n\}$ *does not satisfy condition* (1) *then the Rademacher-series*

$$\sum_{k=0}^{\infty} a_k r_k(x) \tag{3}$$

is not $|C, 1|$*-summable almost everywhere in* $(0, 1)$.

Theorems A and B imply the following statement:

Let $\{a_n\}$ be a given coefficient-sequence. Then there are two cases. Either series (2) is $|C, 1|$-summable for any orthonormal system $\{\varphi_n(x)\}$ on (a, b) almost every-

Received February 17, 1984.

where in (a, b) or the Rademacher-series (3) is not $|C, 1|$-summable almost everywhere in $(0, 1)$.

Later F. Móricz [3] established similar results in the case of absolute Riesz-summability. Very recently H. Schwinn [5] proved an analogous theorem for Euler-means.

In the present paper we shall prove some theorems of this type for an arbitrary regular summability method T. Moreover we give a necessary and sufficient coefficient-condition in order that series (2) for any orthonormal system $\{\varphi_n(x)\}$ be absolutely T-summable (i.e. $|T|$-summable) almost everywhere in the domain of orthogonality.

2. Let $T=(t_{i,n})_{i,n=0}^{\infty}$ be a regular Toeplitz-matrix satisfying the usual conditions:

1. $\lim_{i\to\infty} t_{i,n} = 0 \quad (n = 0, 1, \ldots),$

2. $\lim_{i\to\infty} \sum_{n=0}^{\infty} t_{i,n} = 1,$

3. $\sum_{n=0}^{\infty} |t_{i,n}| \leqq K \quad (<\infty) \quad (i = 0, 1, \ldots).$

Let

$$t_i(a, \varphi; x) = \sum_{n=0}^{\infty} t_{i,n} s_n(x) \quad (i = 0, 1, \ldots), \quad t_{-1}(a, \varphi; x) \equiv 0, \tag{4}$$

where $s_n(x)$ denotes the nth partial sum of (2). Series (2) at a point x_0 is said to be $|T|$-summable if series (4) for each i at x_0 converges and

$$\sum_{i=0}^{\infty} |t_i(a, \varphi; x_0) - t_{i-1}(a, \varphi; x_0)| < \infty.$$

It is clear that the $|T|$-summability of series (2) at x_0 implies the existence of the limit of $t_i(a, \varphi; x_0)$ as $i\to\infty$, i.e. series (2) is also T-summable at x_0.

Let us define the terms $T_{i,k}$ as follows:

$$T_{i,k} = \sum_{n=k}^{\infty} t_{i,n} \quad (i, k = 0, 1, \ldots) \quad \text{and} \quad T_{-1,k} = 0 \quad (k = 0, 1, \ldots).$$

Henceforth let $\{\varphi_n(x)\}$ denote an arbitrary orthonormal system on the σ-finite measure space $(X, \quad, \mu)$. It is clear that if the matrix T is row-finite then

$$t_i(a, \varphi; x) = \sum_{n=0}^{\infty} t_{i,n}(a_0\varphi_0(x) + \ldots + a_n\varphi_n(x)) = \tag{5}$$

$$= \sum_{k=0}^{\infty} T_{i,k} a_k \varphi_k(x) \quad (i = 0, 1, \ldots)$$

holds true at any point x where each function $\varphi_k(x)$ has finite value, i.e. the equality in (5) holds true on X μ-almost everywhere.

If the matrix T is not row-finite, but the sequence $\{a_n\}\in l^2$, then it is easy to show that the equality in (5) also holds true on X μ-almost everywhere. Indeed, if the series on the left-hand side of (5) converges on X μ-almost everywhere to a function $F_i(x)$ and the series on the right-hand side of (5) converges in the metric $L^2(X, \mathscr{A}, \mu)$ to a function $G_i(x)\in L^2(X, \mathscr{A}, \mu)$, i.e.

$$\lim_{N\to\infty} \int_X \Big(\sum_{k=0}^{N} T_{i,k} a_k \varphi_k(x) - G_i(x)\Big)^2 d\mu = 0,$$

then the equality $F_i(x)=G_i(x)$ holds on X μ-almost everywhere $(i=0, 1, \ldots)$.

We prove the following theorems:

Theorem 1. *If T is a row-finite matrix then condition*

$$\sum_{i=0}^{\infty} \Big\{\sum_{k=0}^{\infty} (T_{i,k}-T_{i-1,k})^2 a_k^2\Big\}^{1/2} < \infty \tag{6}$$

implies that series (2) *is* $|T|$*-summable on* X μ*-almost everywhere. If* T *is not a row-finite matrix then* (6) *and* $\{a_n\}\in l^2$ *together imply the* $|T|$*-summability of series* (2) *on* X μ*-almost everywhere.*

Theorem 2. *If*

$$\sum_{i=0}^{\infty} |T_{i,k}-T_{i,k-1}|\,|a_k| < \infty \quad (k = 0, 1, \ldots) \tag{7}$$

and (6) *does not hold, then the Rademacher-series* (3) *is not* $|T|$*-summable almost everywhere in* (0, 1).

Theorem 3. *If the coefficient-sequence* $\{a_k\}$ *does not satisfy condition* (6) *then there exists an orthonormal system* $\{\varphi_n(x)\}$ *such that series* (2) *is not* $|T|$*-summable almost everywhere in* (0, 1).

Remarks. I. It is clear that if the matrix T satisfies the following conditions

$$\sum_{i=0}^{\infty} |T_{i,k}-T_{i-1,k}| < \infty \quad (k = 0, 1, \ldots) \tag{8}$$

then (7) holds true for any coefficient-sequence $\{a_n\}$. An easy calculation shows that the methods of summation $(C, \alpha>0)$ and Riesz satisfy (8).

II. Theorems 1 and 2 imply the cited theorems of P. Billard and F. Móricz; moreover they include the results concerning the $|C, \alpha\geq 1/2|$-summability of the first author [2], and the theorems of H. Schwinn [5] published very recently in connection with Euler summability. All of these assertions can be shown by elementary calculations.

III. By Theorem 3 condition (6) is always necessary in order that series (2) for any orthonormal system $\{\varphi_n(x)\}$ should be $|T|$-summable almost everywhere in the domain of orthogonality.

3. Proofs. Proof of Theorem 1. Let $E\in\mathcal{A}$ with $\mu(E)<\infty$. Then, by (5), we have

$$\sum_{i=0}^{\infty}\int_E |t_i(a,\varphi;x)-t_{i-1}(a,\varphi;x)|\,d\mu \leqq$$

$$\leqq \{\mu(E)\}^{1/2}\sum_{i=0}^{\infty}\Big\{\int_E (t_i(a,\varphi;x)-t_{i-1}(a,\varphi;x))^2\,d\mu\Big\}^{1/2} \leqq$$

$$\leqq \{\mu(E)\}^{1/2}\sum_{i=0}^{\infty}\Big\{\int_X (t_i(a,\varphi;x)-t_{i-1}(a,\varphi;x))^2\,d\mu\Big\}^{1/2} \leqq$$

$$\leqq \{\mu(E)\}^{1/2}\sum_{i=0}^{\infty}\Big\{\sum_{k=0}^{\infty}(T_{i,k}-T_{i-1,k})^2 a_k^2\Big\}^{1/2},$$

which implies that the series

$$\sum_{i=0}^{\infty}|t_i(a,\varphi;x)-t_{i-1}(a,\varphi;x)| \tag{9}$$

converges on E μ-almost everywhere. By the assumption the measure space $(X,\mathcal{A},\mu)$ is σ-finite, so it also follows that series (9) converges on X μ-almost everywhere, that is, series (2) is $|T|$-summable on X μ-almost everywhere, as desired.

Proof of Theorem 2. We distinguish two cases. If $\{a_n\}\notin l^2$ then by a well-known theorem of A. Zygmund [8] the Rademacher-series (3) is not T-summable almost everywhere in $(0,1)$, and consequently it is not $|T|$-summable almost everywhere in $(0,1)$. In this case our theorem is already proved.

Next let us assume that $\{a_n\}\in l^2$. In this case we need a slightly modified version of a well-known theorem of Orlicz [4]. We formulate it as a lemma.

Lemma. *For any Lebesgue-measurable set $E(\subseteqq(0,1))$ there exist a positive number $K=K(E)$ and a natural number $k_0=k_0(E)$ such that if $\{a_n\}\in l^2$ and $k_1\geqq k_0$ then*

$$K(E)(\mathrm{mes}\,E)\Big\{\sum_{k=k_1}^{\infty}a_k^2\Big\}^{1/2}\leqq\int_E\Big|\sum_{k=k_1}^{\infty}a_k r_k(t)\Big|\,dt$$

holds true.

Returning to the proof of Theorem 2, if now we assume the contrary of the statement of Theorem 2, that is, that series (3) is $|T|$-summable on a set $E(\subseteqq(0,1))$

of positive measure, then there exist a positive number M and a set $F(\subseteqq E)$ of positive Lebesgue measure such that

$$\sum_{i=0}^{\infty} |t_i(a, \varphi; x) - t_{i-1}(a, \varphi; x)| \leqq M \tag{10}$$

holds for any $x \in F$.

Then, by Lemma, there exists a natural number $k_0 = k_0(F)$ such that

$$\int_F \Big| \sum_{k=k_0}^{\infty} (T_{i,k} - T_{i-1,k}) a_k r_k(x) \Big| \, dx \geqq \tag{11}$$
$$\geqq K(F)(\text{mes } F) \Big\{ \sum_{k=k_0}^{\infty} (T_{i,k} - T_{i-1,k})^2 a_k^2 \Big\}^{1/2}.$$

Moreover, by (7), we have

$$\sum_{i=0}^{\infty} \Big| \sum_{k=0}^{k_0-1} (T_{i,k} - T_{i-1,k}) a_k r_k(x) \Big| \leqq \sum_{k=0}^{k_0-1} M(k), \tag{12}$$

where

$$M(k) := \sum_{i=0}^{\infty} |T_{i,k} - T_{i-1,k}| \cdot |a_k| \quad (k = 0, 1, \ldots).$$

Now, using (5), (7), (10), (11) and (12) yield

$$K(F)(\text{mes } F) \sum_{i=0}^{\infty} \Big\{ \sum_{k=k_0}^{\infty} (T_{i,k} - T_{i-1,k})^2 a_k^2 \Big\}^{1/2} \leqq \tag{13}$$
$$\leqq (\text{mes } F)\Big(M + \sum_{k=0}^{k_0-1} M(k)\Big) < \infty,$$

furthermore, using (7) once more, we get

$$\sum_{i=0}^{\infty} \Big\{ \sum_{k=0}^{k_0-1} (T_{i,k} - T_{i-1,k})^2 a_k^2 \Big\}^{1/2} \leqq \sum_{k=0}^{k_0-1} M(k) < \infty. \tag{14}$$

Estimations (13) and (14) imply that (6) holds true, which is a contradiction, and this proves Theorem 2.

Proof of Theorem 3. We distinguish two cases again.

If (7) holds true for each k then, by Theorem 2, the Rademacher-series (3) is not $|T|$-summable almost everywhere in $(0, 1)$.

If (7) does not hold for a certain natural number k_0, that is,

$$\sum_{i=0}^{\infty} |T_{i,k_0} - T_{i-1,k_0}| \cdot |a_{k_0}| = \infty,$$

then we define a special orthonormal system $\{\psi_n(x)\}$ as follows. Let

$$\psi_{k_0}(x) = \begin{cases} \sqrt{2}, & x \in (0, 1/2), \\ 0, & x \in (1/2, 1), \end{cases}$$

furthermore let us choose the functions $\psi_k(x)$ $\{k=0, 1, \ldots; k\neq k_0\}$ such that they are zero on $(0, 1/2)$ and form an orthonormal system on the interval $(1/2, 1)$. Then the system $\{\psi_n(x)\}_0^\infty$ is orthonormal on $(0, 1)$. For this system we obviously have

$$\int_0^1 \Big(\sum_{i=0}^{\infty} |t_i(a, \psi; x)-t_{i-1}(a, \psi; x)|\Big)\, dx \geqq$$

$$\geqq \int_0^{1/2} \Big(\sum_{i=0}^{\infty} |t_i(a, \psi; x)-t_{i-1}(a, \psi; x)|\Big)\, dx =$$

$$= (\sqrt{2}/2)\, |a_{k_0}| \sum_{i=0}^{\infty} |T_{i,k_0}-T_{i-1,k_0}| = \infty,$$

whence

$$\|a; T\| := \sup_{\{\varphi_k\}} \int_0^1 \Big(\sum_{i=0}^{\infty} |t_i(a, \varphi; x)-t_{i-1}(a, \varphi; x)|\Big)\, dx = \infty \tag{15}$$

follows, where the supremum is taken for all orthonormal systems $\{\varphi_n(x)\}$ on (a, b). On account of a theorem of the second author [7] statement (15) implies the existance of an orthonormal system $\{\varphi_n(x)\}$ for which series (2) is not $|T|$-summable almost everywhere in $(0, 1)$.

This completes the proof.

References

[1] P. Billard, Sur la sommabilité absolue des séries de fonctions orthogonales, *Bull. Sci. Math.*, **85** (1961), 29—33.

[2] L. Leindler, Über die absolute Summierbarkeit der Orthogonalreihen, *Acta Sci. Math.*, **22** (1961), 243—268.

[3] F. Móricz, Über die Rieszsche Summation der Orthogonalreihen, *Acta Sci. Math.*, **23** (1962), 92—95.

[4] W. Orlicz, Beiträge zur Theorie der Orthogonalentwicklungen, *Studia Math.*, **6** (1936), 20—38.

[5] H. Schwinn, Absolute Summability of Orthogonal Series by Euler-Means, *Analysis*, **2** (1982), 219—230.

[6] K. Tandori, Über die orthogonalen Funktionen. IX (Absolute Summation), *Acta Sci. Math.*, **21** (1960), 292—299.

[7] K. Tandori, Über die absolute Summierbarkeit der Orthogonalreihen, *Acta Math. Acad. Sci. Hungar.*, **22** (1971), 215—226.

[8] A. Zygmund, On the convergence of lacunary trigonometric series, *Fund. Math.*, **16** (1930), 90—107.

BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

Acta Sci. Math., 50 (1986), 105—124

Uniform laws of the iterated logarithm for Lipschitz classes of functions

SOMPONG DHOMPONGSA*)

1. Introduction

Let $\{[0,1], \mathscr{F}, P\}$ be the unit interval with Lebesgue measurability and Lebesgue measure P. For $1/2<\alpha$, let Λ_α be the class of real-valued functions f on $[0,1]$ with $f(0)=f(1)$, $\int_0^1 f(x)\,dx=0$ and satisfying a Lipschitz condition

$$|f(x)-f(y)| \leqq |x-y|^\alpha, \quad 0 \leqq x, y \leqq 1.$$

Extend the functions of Λ_α with period 1.

The purpose of this paper is to prove the following two theorems.

Theorem 1. *Let $\{n_k, k\geqq 1\}$ be a sequence of real numbers satisfying*

$$n_{k+1}/n_k \geqq 1+c/k^\delta \quad (c>0) \tag{1.1}$$

for some $0<\delta<1/2$. Then

$$\limsup_{N\to\infty} \sup_{f\in\Lambda_\alpha} \Big|\sum_{k\leqq N} f(n_k x)\Big|/(N\log\log N)^{1/2} \leqq C \quad a.s.$$

(with respect to the Lebesgue measure on $[0,1]$). The constant C depends only on α and δ.

We say that a sequence $\{n_k\}$ of integers satisfies condition B_2 if there is a constant C such that the number of solutions of the equation $n_k \pm n_l = v$ does not exceed C for any $v\geqq 0$.

Theorem 2. *Let $\{n_k, k\geqq 1\}$ be a sequence of integers satisfying condition B_2 and*

$$n_{k+1}/n_k \geqq 1+c/k^\delta \quad (c>0) \tag{1.2}$$

Received January 11, 1984.

*) Supported by the Thai Government.

with $1/2 \leqq \delta \leqq 1$. *Then for each* α *with* $1/2+\delta/2<\alpha$,

$$\limsup_{N\to\infty} \sup_{f\in A_\alpha} \Big|\sum_{k\leqq N} f(n_k x)\Big|/(N\log\log N)^{1/2} \leqq C$$

for almost all $x\in[0, 1]$, *where* C *is a constant depending on* α, δ *and the constant in the* B_2 *condition of the sequence* $\{n_k, k\geqq 1\}$.

The results in Theorem 1 improve upon Theorem 3.2 of Kaufman and Philipp [10] who, instead of (1.1), assume the more restrictive condition

$$n_{k+1}/n_k \geqq q > 1.$$

2. Proof of Theorem 1

In the course of the proof of Theorem 1 we shall prove the following two propositions.

Proposition 1. *Let* $\{n_k, k\geqq 1\}$ *be as in Theorem* 1. *Then there exist positive constants* A *and* C_1 *such that*

$$P\Big\{\Big|\sum_{k=1}^{N} \exp(2\pi i n_k x)\Big| \geqq AR(N\log\log N)^{1/2}\Big\} \leqq C_1 \exp(-10R\log\log N)$$

for all $R\geqq 1$, $N\geqq 1$.

Proposition 2. *Let* $\{n_k, k\geqq 1\}$ *be as in Theorem* 1. *Then*

$$P\Big\{\max_{1\leqq m<N}\Big|\sum_{k=N+1}^{N+m} \exp(2\pi i n_k x)\Big| \geqq AR(N\log\log N)^{1/2}\Big\} \leqq C_1 \exp(-10R\log\log N)$$

for all $R\geqq 1$, $N\geqq 1$. *Here* A *and* C_1 *are as in Proposition* 1.

Before we prove these propositions we shall deduce Theorem 1 from them. We need to introduce some notations. For integers h, N, $H\geqq 1$, we put

$$F(H, N, h) = \Big|\sum_{k=H+1}^{H+N} \exp(2\pi i h n_k x)\Big|. \tag{2.1}$$

Then

$$F(0, N, h) \leqq F(0, 2^n, h) + \max_{1\leqq m<2^n} F(2^n, m, h) \tag{2.2}$$

where $n=[\log N/\log 2]$. Here $[x]$ denotes the largest integer not exceeding x for any real number x. Put

$$\varphi(N) = (N\log\log N)^{1/2} \tag{2.3}$$

and define the events (here and throughout $\log^+ x = \log(\max(e, x))$)

$$G(n, h) = \{F(0, 2^n, h) \geqq A \log^+ |h| \varphi(2^n)\}; \quad G_n = \bigcup_{|h| \geqq 1} G(n, h);$$

$$H(n, h) = \Big\{\max_{1 \leqq m < 2^n} F(2^n, m, h) \geqq A \log^+ |h| \varphi(2^n)\Big\}; \quad H_n = \bigcup_{|h| \geqq 1} H(n, h),$$

where A is the constant appearing in Proposition 1.

We now show that with probability 1 only finitely many G_n and H_n occur. In fact, by Proposition 1,

$$P\{G(n, h)\} \ll \exp(-2 \log^+ |h| \log n).$$

Thus

$$P(G_n) \ll \sum_{|h| \geqq 1} \exp(-2 \log^+ |h| \log n) \ll n^{-2}.$$

Also, by Proposition 2,

$$P\{H(n, h)\} \ll \exp(-2 \log^+ |h| \log n).$$

Thus

$$P(H_n) \ll \sum_{|h| \geqq 1} \exp(-2 \log^+ |h| \log n) \ll n^{-2}.$$

Now we can conclude that

$$F(0, N, h) \ll \log^+ |h| \varphi(N) \quad \text{a.s.} \tag{2.4}$$

for all $|h| \geqq 1$ by (2.2).

In [10], Kaufman and Philipp showed that if $f \in \Lambda_\alpha$ $(\alpha > 1/2)$, then the coefficients a_h of the Fourier series of f

$$f(x) = \sum_{|h| \geqq 1} a_h \exp(2\pi i h x) \tag{2.5}$$

satisfies

$$\sum_{|h| \geqq N} a_h \exp(2\pi i h x) \ll N^{-1/2} \tag{2.6}$$

uniformly in x, and

$$\sum_{|h| \geqq 1} |a_h|^2 |h| (\log^+ |h|)^4 \ll 1. \tag{2.7}$$

In fact, (2.7) can be replaced by

$$\sum_{|h| \geqq 1} |a_h|^2 |h|^{1+\varepsilon} \ll 1 \tag{2.8}$$

for any ε with $0 < \varepsilon < 2\alpha - 1$ since in [15], formula (3) on page 136, we have

$$\sum_{h=2^{\nu-1}-1}^{2^\nu} |a_h|^2 < 2C 2^{-2\alpha\nu}$$

for an absolute constant C. Thus

$$(2.9)\qquad \sum_{h=2^{v-1}-1}^{2^v} |a_h|^2 h^{1-\varepsilon} \leqq 2^{(1+\varepsilon)v} \sum_{h=2^{v-1}-1}^{2^v} |a_h|^2 \leqq 2C2^{-(2\alpha-1-\varepsilon)v}.$$

Since $\varepsilon < 2\alpha - 1$, (2.8) follows from (2.9).

Now for each $f \in A_\alpha$ $(\alpha > 1/2)$, we have if $0 < \varepsilon' < \varepsilon < 2\alpha - 1$

(2.10)

$$\Big|\sum_{k\leqq N} f(n_k x)\Big| \leqq \Big|\sum_{1\leqq|h|\leqq N/2} a_h \sum_{k\leqq N} \exp(2\pi i n_k hx)\Big| + \Big|\sum_{k\leqq N}\sum_{|h|>N/2} a_h \exp(2\pi i n_k hx)\Big| \ll$$
$$\ll \Big(\sum_{1\leqq|h|\leqq N/2} |a_h|^2 |h|^{1+\varepsilon}\Big)^{1/2} \Big(\sum_{1\leqq|h|\leqq N/2} |h|^{-1-\varepsilon} F^2(0, N, h)\Big)^{1/2} + NN^{-1/2} \ll$$
$$\ll 1\Big(\sum_{1\leqq|h|\leqq N/2} |h|^{-1-\varepsilon}|h|^{\varepsilon'}\Big)^{1/2} \varphi(N) + N^{1/2} \ll \varphi(N)$$

by (2.5), (2.1), (2.6), (2.8), and (2.4).

2.1 Proof of Proposition 1. (The proof is based on the idea of the proof of Proposition 4.2.1 of Philipp [11].)

Lemma 1. *For* $1 \leqq j < k$ *we have* $n_j/n_k \leqq 2^{-c(k-j)/k^\delta}$.

The proof is very simple (cf. [1], p. 211).

Choose ε so that

$$(2.11)\qquad \delta < \varepsilon/(1+\varepsilon) < 1/2.$$

We now divide $\mathbf{Z}^+$ into blocks (without gaps) such that

$$H_1 < I_1 < H_2 < I_2 < \ldots < H_j < I_j < \ldots \quad \text{(say)},$$
$$\operatorname{card}(H_j) = \operatorname{card}(I_j) = [j^\varepsilon].$$

Let

(2.12) c_j be the smallest element of H_j, and let d_j be the largest element of H_j.

For $v \in \mathbf{Z}^+$, we write

$$(2.13)\qquad m_v = [\varrho_v + 5\log v]$$

where $2^{\varrho_v} \leqq n_v < 2^{\varrho_v+1}$. Put

$$(2.14)\qquad \psi_v(x) = \cos 2\pi n_v(k/2^{m_v}) \quad \text{if} \quad x \in [k2^{-m_v}, (k+1)2^{-m_v})$$
$$(k = 0, 1, \ldots, 2^{m_v}-1).$$

Now write

$$(2.15)\quad T_j = \sum_{v\in H_j} \cos 2\pi n_v x, \quad D_j = \sum_{v\in H_j} \psi_v(x), \quad \bar{D}_j = D_j - E(D_j | D_1, \ldots, D_{j-1}).$$

(The definition of D_j was introduced by Berkes [1].)

Lemma 2. $\|T_j - D_j\|_\infty \ll C_j^{-3}$.

Proof. We have

$$|\cos 2\pi n_v x - \psi_v(x)| \leqq 2\pi n_v 2^{-m_v} \ll n_v 2^{-e_v} v^{-4} \ll v^{-4}$$

by (2.13). Thus

$$\|T_j - D_j\|_\infty \ll \sum_{v \in H_j} v^{-4} \ll C_j^{-3}.$$

by (2.12).

Lemma 3. $E(T_j | D_1, \ldots, D_{j-1}) \ll j^{-2}$ *a.s.*

Proof. Let

$$A = [k2^{-m_{d_{j-1}}}, (k+1)2^{-m_{d_{j-1}}})$$

for any $k=0, 1, \ldots, 2^{m_{d_{j-1}}}$. Then, putting $m_v' = 2\pi n_v / 2^{m_{d_{j-1}}}$, we have

$$P^{-1}(A) \int_A T_j = \int_k^{k+1} \sum_{v \in H_j} \cos m_v' x \ll$$

$$\ll \sum_{v \in H_j} 1/m_v' \ll j^\varepsilon / m_{c_j}' \ll j^\varepsilon 2^{m_{d_{j-1}}} / n_{c_j} \ll$$

$$\ll j^\varepsilon d_{j-1}^5 (n_{d_{j-1}} / n_{c_j}) \ll j^\varepsilon d_{j-1}^5 2^{-cj^\varepsilon / j^{(1+\varepsilon)\delta}} \ll j^{\varepsilon + 5(1+\varepsilon)} 2^{-cj^{\varepsilon - (1+\varepsilon)\delta}} \ll j^{-2}$$

by Lemma 1 and (2.12).

Lemma 4. $E(T_j^2 | D_1, \ldots, D_{j-1}) \ll j^\varepsilon$ *a.s.*

Proof. Let A be as in the proof of Lemma 3.

$$P^{-1}(A) \int_A T^2 = \int_k^{k+1} \Big(\sum_{v \in H_j} \cos m_v' x\Big)^2 \ll$$

$$\ll j^\varepsilon + \Big|\int_k^{k+1} \sum_{\mu < v \in H_j} \cos m_\mu' x \cos m_v' x\Big| + \Big|\int_k^{k+1} \sum_{v \in H_j} \cos 2m_v' x\Big| \ll$$

$$\ll j^\varepsilon + j^\varepsilon / m_{c_j}' + j^\varepsilon \sum_{v \in H_j} 1/(m_{v+1}' - m_v') \ll j^\varepsilon + j^\varepsilon d_j^\delta \sum_{v \in H_j} 1/m_v' \ll$$

$$\ll j^\varepsilon + j^{\varepsilon + (1+\varepsilon)\delta} j^{\varepsilon + 5(1+\varepsilon)} 2^{-cj^{\varepsilon - (1+\varepsilon)\delta}} \ll j^\varepsilon \quad \text{a.s.}$$

by (2.13) and (2.12).

Lemma 5. (1) $E(D_j | D_1, \ldots, D_{j-1}) \ll j^{-2}$ *a.s.*
(2) $E(\bar{D}_j^2 | D_1, \ldots, D_{j-1}) \ll j^\varepsilon$ *a.s.*

Proof. (1) follows from Lemma 2 and (2.13). (2) follows from Lemma 4 and the following computation: Since $\mathrm{Var}(X|\mathscr{A}) \leqq E(X^2|\mathscr{A})$, we have, by (2.15),

$$E(\bar{D}_j^2|D_1, \ldots, D_{j-1}) \leqq E(D_j^2|D_1, \ldots, D_{j-1}) + E^2(D_j|D_1, \ldots, D_{j-1}) \ll$$
$$\ll E(|D_j^2 - T_j^2|\,|D_1, \ldots, D_{j-1}) + E(T_j^2|D_1, \ldots, D_{j-1}) \ll$$
$$\ll j^{\varepsilon}\|T_j - D_j\|_\infty + E(T_j^2|D_1, \ldots, D_{j-1}) \ll j^{\varepsilon-3(1+\varepsilon)} + E(T_j^2|D_1, \ldots, D_{j-1}).$$

Lemma 6. *Let $B(\geqq 1)$ be the constant implied by $\ll$ in Lemma 5(2). Then we have as $M \to \infty$*

$$P\Big\{\Big|\sum_{j\leqq M} \bar{D}_j\Big| > 16RB(M^{\varepsilon+1}\log\log M^{\varepsilon+1})^{1/2}\Big\} \ll \exp(-12RB\log\log M^{\varepsilon+1}).$$

Proof. Put

$$u_n = \sum_{j\leqq n} \bar{D}_j, \quad n \leqq M,$$
$$= u_M, \quad n > M;$$
$$s_n^2 = \sum_{j\leqq n} E(\bar{D}_j^2|D_1, \ldots, D_{j-1}), \quad n \leqq M,$$
$$= s_M^2, \qquad n > M;$$
$$c = M^{\varepsilon}, \quad \lambda = 2(\log\log M^{\varepsilon+1})^{1/2} M^{-(\varepsilon+1)/2}, \quad K = 4RBM^{\varepsilon+1};$$
$$T_n = \exp\big(\lambda u_n - (1/2)\lambda^2(1 + (1/2)\lambda c)s_n^2\big).$$

Thus as in the proof of Lemma 4.2.9 of Philipp [11],

$$P\Big\{\sum_{j\leqq M} \bar{D}_j > 8RB(M^{\varepsilon+1}\log\log M^{\varepsilon+1})^{1/2}\Big\} \leqq P\Big\{\sup_{n\geqq 0} u_n > \lambda K\Big\} \leqq$$
$$\leqq P\Big\{\sup_{n\geqq 0} T_n > \exp(\lambda^2 K - \lambda^2 BM^{\varepsilon+1})\Big\} \leqq \exp(\lambda^2 BM^{\varepsilon+1} - \lambda^2 K) \leqq$$
$$\leqq \exp(-12RB\log\log M^{\varepsilon+1}).$$

Lemma 7. *There is a positive constant A_1 such that*

$$P\Big\{\Big|\sum_{j\leqq M} T_j\Big| \geqq A_1 R(M^{\varepsilon+1}\log\log M^{\varepsilon+1})^{1/2}\Big\} \ll \exp(-12R\log\log M^{\varepsilon+1}).$$

Proof. This lemma is a consequence of Lemmas 2, 5(1) and 6 together with the following equality:

$$\sum_{j\leqq M} T_j = \sum_{j\leqq M} (T_j - D_j) + \sum_{j\leqq M} (D_j - \bar{D}_j) + \sum_{j\leqq M} \bar{D}_j.$$

Similarly, we can prove that

$$P\Big\{\Big|\sum_{j\leqq M} T_j'\Big| \geqq A_1 R(M^{\varepsilon+1}\log\log M^{\varepsilon+1})^{1/2}\Big\} \ll \exp(-12R\log\log M^{\varepsilon+1})$$

where

$$T_j' = \sum_{v \in I_j} \cos 2\pi n_v x. \tag{2.16}$$

Hence if $N \in H_{M+1}$, then, by (2.15), (2.16) and Lemma 7,

$$P\Big\{\Big|\sum_{k \le N} \cos 2\pi n_k x\Big| \ge 3A_1 R(N \log\log N)^{1/2}\Big\} \le P\Big\{\Big|\sum_{j \le M} T_j\Big| \ge A_1 R(N \log\log N)^{1/2}\Big\} +$$
$$+ P\Big\{\Big|\sum_{j \le M} T_j'\Big| \ge A_1 R(N \log\log N)^{1/2}\Big\} +$$
$$+ P\Big\{\Big|\sum_{v=C_{M+1}}^{N} \cos 2\pi n_v x\Big| \ge A_1 R(N \log\log N)^{1/2}\Big\} \ll$$
$$\ll \exp(-11R \log\log N) + \exp(-11R \log\log N) + 0 \ll \exp(-11R \log\log N).$$

We have used the fact that for all large N, $N - C_{M+1} \le (M+1)^{\varepsilon} < N^{1/2}$ since $\varepsilon < (\varepsilon+1)/2$ by (2.11).

The case where $N \in I_M$ can be proved in the same way. Thus, in general, we have

$$P\Big\{\Big|\sum_{k \le N} \cos 2\pi n_k x\Big| \ge 3A_1 R(N \log\log N)^{1/2}\Big\} \ll \exp(-11R \log\log N).$$

Similarly, we can prove that

$$P\Big\{\Big|\sum_{k \le N} \sin 2\pi n_k x\Big| \ge 3A_1 R(N \log\log N)^{1/2}\Big\} \ll \exp(-11R \log\log N).$$

Hence Proposition 1 is proved.

2.2 Proof of Proposition 2.

Lemma 8. *Put*

$$Z_k = \sum_{v=Q+1}^{Q+k} \psi_v(x).$$

Then we have for any $Q \ge 0$ *and any real number* t

$$P\Big\{\max_{1 \le k \le N} |Z_k| > t\Big\} \le 2P\Big\{|Z_n| > t - 4\sqrt{N}\Big\}$$

provided that $N \ge N_0$, $Q \le N^{1+\gamma}$ *where* γ *is a positive constant such that* $(1+\gamma)\delta < 1/2$.

The proof of this lemma is exactly the same as the one in [1], pp. 214—216. Note that in the proof the n_k's do not have to be integers.

Lemma 9. *We have for any* $0 \le Q \le N^{1+\gamma}$, $N \ge N_0$, *and any real number* $t \ge 3\sqrt{N}$,

$$P\Big\{\max_{1 \le k \le N} \Big|\sum_{v=Q+1}^{Q+k} \cos 2\pi n_v x\Big| > 3t\Big\} \le 2P\Big\{\Big|\sum_{v=Q+1}^{Q+N} \cos 2\pi n_v x\Big| > t - 2\sqrt{N}\Big\}.$$

Proof. By the definition of ψ_v in (2.14) we have, as before,

$$\|\cos 2\pi n_v x - \psi_v(x)\|_\infty \ll v^{-4}.$$

Thus

$$\max_{1\leq k\leq N} \Big\| \sum_{v=Q+1}^{Q+k} \cos 2\pi n_v x - \psi_v(x)\Big\|_\infty \ll 1.$$

Hence, if t is large enough,

$$P\Big\{\max_{1\leq k\leq N} \Big| \sum_{v=Q+1}^{Q+k} \cos 2\pi n_v x - \psi_v(x)\Big| > t\Big\} = 0.$$

Now Lemma 9 follows from Lemma 8.

Lemma 10. *As* $N\to\infty$,

$$P\Big\{\Big|\sum_{k=N+1}^{2N-1} \cos 2\pi n_k x\Big| \geqq 8A_1 R(N\log\log N)^{1/2}\Big\} \ll \exp(-10R\log\log N).$$

Proof. The probability in question does not exceed the probability

$$P\Big\{\Big|\sum_{k=1}^{2N-1} \cos 2\pi n_k x\Big| \geqq 5A_1 R(N\log\log N)^{1/2}\Big\} +$$

$$+P\Big\{\Big|\sum_{k=1}^{N} \cos 2\pi n_k x\Big| \geqq 3A_1 R(N\log\log N)^{1/2}\Big\} \ll$$

$$\ll P\Big\{\Big|\sum_{k=1}^{2N-1} \cos 2\pi n_k x\Big| \geqq 3A_1 R(2N\log\log 2N)^{1/2}\Big\} + \exp(-10R\log\log N)$$

by Proposition 1.

By Lemmas 9 and 10 we can say that

$$P\Big\{\max_{1\leq m<N} \Big|\sum_{k=N+1}^{N+m} \cos 2\pi n_k x\Big| \geqq 30A_1 R(N\log\log N)^{1/2}\Big\} \leqq$$

$$\leqq 2P\Big\{\Big|\sum_{k=N+1}^{2N-1} \cos 2\pi n_k x\Big| \geqq 10A_1 R(N\log\log N)^{1/2} - 2\sqrt{N}\Big\} \ll$$

$$\ll P\Big\{\Big|\sum_{k=N+1}^{2N-1} \cos 2\pi n_k x\Big| \geqq 8A_1 R(N\log\log N)^{1/2}\Big\} \ll \exp(-10R\log\log N).$$

Similarly, we can show that

$$P\Big\{\max_{1\leq m<N} \Big|\sum_{k=N+1}^{N+m} \sin 2\pi n_k x\Big| \geqq 30A_1 R(N\log\log N)^{1/2}\Big\} \ll \exp(-10R\log\log N).$$

Thus Proposition 2 follows. Also we can choose A and C_1 so large that both propositions will apply.

3. Proof of Theorem 2

We assume that $\alpha \leqq 1$ (the case $\alpha > 1$ is trivial). We choose ε so that

$$\varepsilon > \delta/(1-\delta) \quad \text{and} \quad 1/2+\varepsilon/2(\varepsilon+1) < \alpha. \tag{3.1}$$

This can be done by choosing ε sufficiently close to $\delta/(1-\delta)$ since $1/2+\delta/2<\alpha$ and since $\delta<1$. Put

$$\gamma = 2/(2\alpha+1). \tag{3.2}$$

Thus, by (3.1) $\varepsilon/2(\varepsilon+1)<(1-\gamma)/\gamma \leqq 1/2$. Now choose ε' and β so that

$$\varepsilon/2(\varepsilon+1) < \varepsilon' < (1-\gamma)/\gamma \tag{3.3}$$

and

$$1/2\varepsilon' > \beta > \gamma/2(1-\gamma). \tag{3.4}$$

Put

$$q = \beta/(\beta-1). \tag{3.5}$$

Thus, by (3.4)

$$-q+2q\varepsilon' < -1. \tag{3.6}$$

As in Section 2 we shall prove the following two propositions.

Proposition 3. *Let $\{n_k,\ k \geqq 1\}$ be as in Theorem 2. Then there exist positive constants A and C_1 such that*

$$P\Big\{\Big|\sum_{k=1}^{N} \exp(2\pi i n_k x)\Big| \geqq AR(N\log\log N)^{1/2}\Big\} \leqq$$
$$\leqq C_1 \exp(-10R\log\log N) + C_1 R^{-1} N^{-1/(\varepsilon+1)}(\log\log N) +$$
$$+ C_1 R^{-2} N^{-1/2(\varepsilon+1)}(\log\log N)$$

for all $R \geqq 1$, $N \geqq 1$, where ε is defined in (3.1).

Proposition 4. *Let $\{n_k,\ k \geqq 1\}$ be as in Theorem 2. Then*

$$P\Big\{\max_{1\leqq m\leqq N}\Big|\sum_{k=N+1}^{N+m} \exp(2\pi i n_k x)\Big| \geqq AR(N\log\log N)^{1/2}\Big\} \leqq$$
$$\leqq C_1 \exp(-10R\log\log N) + C_1 R^{-1} N^{-1/(\varepsilon+1)}(\log\log N) +$$
$$+ C_1 R^{-2} N^{-1/2(\varepsilon+1)}(\log\log N)$$

for all $R \geqq 1$, $N \geqq 1$. Here A and C_1 are the same as in Proposition 3.

The proofs of these two propositions will be given in Section 3.1 and 3.2.

To apply the propositions we need to define some events. For integers h, $n \geqq 0$, let

$$G(n, h) = \{F(0, 2^n, h) \geqq A|h|^{2\varepsilon'}\varphi(2^n)\};$$

$$G_n = \bigcup_{1 \leqq |h| \leqq 2^n} G(n, h);$$

$$H(n, h) = \{\max_{1 \leqq m < 2^n} F(2^n, m, h) \geqq A|h|^{2\varepsilon'}\varphi(2^n)\};$$

$$H_n = \bigcup_{1 \leqq |h| \leqq 2^n} H(n, h)$$

by using definitions in (2.1), (3.4) and (2.3). Here A is the constant appearing in Proposition 3.

Taking Propositions 3 and 4 for granted we see that

$$P(G_n) \ll \sum_{1 \leqq |h| \leqq 2^n} \exp(-2|h|^{2\varepsilon'} \log n) + (|h|^{-2\varepsilon'} 2^{(-1/(\varepsilon+1))n} + |h|^{-4\varepsilon'} 2^{(-1/2(\varepsilon+1))n}) \log n) \ll$$

$$\ll n^{-2} + 2^{(1-2\varepsilon'-1/(\varepsilon+1))n} \log n + 2^{(1-4\varepsilon'-1/2(\varepsilon+1))n} \log n;$$

$$P(H_n) \ll n^{-2} + 2^{(1-2\varepsilon'-1/(\varepsilon+1))n} \log n + 2^{(1-4\varepsilon'-1/2(\varepsilon+1))n} \log n.$$

Since $1-2\varepsilon' < 1/(\varepsilon+1)$ and $1-4\varepsilon' < 1/2(\varepsilon+1)$ by (3.3) and (3.1), we have

$$\sum_{n \geqq 0} P(G_n) < \infty, \quad \sum_{n \geqq 0} P(H_n) < \infty.$$

Thus by Borel—Cantelli Lemma,

$$F(0, N, h) \ll |h|^{2\varepsilon'} \varphi(N) \quad \text{a.s.}$$

for all $1 \leqq |h| \leqq N/2$.

Now for each $f \in A_\alpha$ $(\alpha > 1/2 + \delta/2)$ we have as in (2.10)

$$(3.7) \quad \Big|\sum_{k \leqq N} f(n_k x)\Big| \leqq \Big(\sum_{1 \leqq |h| \leqq N/2} |a_h|^\beta |h|^\beta\Big)^{1/\beta} \Big(\sum_{1 \leqq |h| \leqq N/2} |h|^{-q} F^q(0, N, h)\Big)^{1/q} + N^{1/2} \ll$$

$$\ll 1\Big(\sum_{|h| \geqq 1} |h|^{-q+2\varepsilon' q}\Big)^{1/q} \varphi(N) + N^{1/2} \ll \varphi(N)$$

by (3.4), (3.5), (3.6) and (2.6).

We have used the fact that, by (3.4) and (3.2), $\beta > \gamma$. Thus the argument in Section 6.32 on page 137 of [15] implies

$$(3.8) \quad \sum_{h=2^{v-1}+1}^{2^v} |a_h|^\beta |h|^\beta \ll 2^{\beta v} 2^{v((1/2)-(\beta/\gamma))} \ll 2^{v(\beta+(1/2)-(\beta/\gamma))}.$$

Since $\beta + (1/2) - (\beta/\gamma) < 0$ by (3.4), (3.8) implies

$$\sum_{|h| \geqq 1} |a_h|^\beta |h|^\beta \ll 1.$$

3.1 **Proof of Proposition 3.** The proof of Proposition 3 is similar to that of Proposition 1. We begin with an estimate for fourth moments.

Lemma 11. *For integers* $M \geqq 0$, $N \geqq 1$, *put* $S_{M,N} = \sum_{j=M+1}^{M+N} \cos 2\pi n_j x$. *Thus*

$$\int_0^1 S_{M,N}^4 \ll N^2$$

where the constant implied by $\ll$ *depends on bound in* B_2*-condition of the sequence* $\{n_k\}$.

This lemma is Lemma (5.4) of [3]. The proof is quite simple. It is strongly based on the assumptions that $n_k \in \mathbf{Z}$ and satisfies B_2-condition.

Lemma 12. *For* $1 \leqq j < k$ *we have* $n_j/n_k \leqq 2^{-c(k-j)/k^\delta}$.

Proof.

$$n_k/n_j \geqq \prod_{i=j}^{k-1} (1+c/i^\delta) > (1+c/k^\delta)^{k-j} > 2^{c(k-j)/k^\delta}$$

by (1.2).

Divide $\mathbf{Z}^+$ into blocks (without gaps) such that

$$H_1 < I_1 < H_2 < I_2 < \ldots < H_j < I_j < \ldots \quad \text{(say)},$$
$$\operatorname{card}(H_j) = [j^\varepsilon] = \operatorname{card}(I_j)$$

where ε is defined in (3.1). We define c_j, d_j, m_v, ϱ_v, ψ_v, T_j, D_j, and $\bar{D}_j$ as in Section 2.1.

The proofs of the following three lemmas are the same as the proofs of Lemma 2, 3 and 4 respectively.

Lemma 13. $\|T_j - D_j\|_\infty \ll C_j^{-3}$.

Lemma 14. $E(T_j | D_1, \ldots, D_{j-1}) \ll j^{-2}$ *a.s.*

Lemma 15. $E(T_j^2 | D_1, \ldots, D_{j-1}) \ll j^\varepsilon$ *a.s.*

Lemma 16. (1) $E(D_j | D_1, \ldots, D_{j-1}) \ll j^{-2}$ *a.s.*
(2) $E(D_j^2 | D_1, \ldots, D_{j-1}) \ll j^\varepsilon$ *a.s.*
(3) $E(\bar{D}_j^2 | D_1, \ldots, D_{j-1}) \ll j^\varepsilon$ *a.s.*

Lemma 16 can be proved as in Lemma 5.

Put

$$t_j = j^{(\varepsilon+1)/2} / 4\sqrt{\log\log j^{\varepsilon+1}}, \tag{3.9}$$

and define random variables

$$D_j^* = D_j I[|D_j| \leqq t_M] \qquad (j \leqq M), \tag{3.10}$$
$$\bar{D}_j^* = D_j^* - E(D_j^* | D_1, \ldots, D_{j-1}) \quad (j \leqq M).$$

Thus

(3.11) $$E(\bar{D}_j^{*2}|D_1, \ldots, D_{j-1}) \ll j^{\varepsilon} \quad \text{a.s.}$$

since

$$E(\bar{D}_j^{*2}|D_1, \ldots, D_{j-1}) \leqq E(D_j^{*2}|D_1, \ldots, D_{j-1}) \leqq E(D_j^2|D_1, \ldots, D_{j-1}) \ll j^{\varepsilon} \quad \text{a.s.}$$

by Lemma 16(2).

Lemma 17. *Let B ($\geqq 2$) be the constant implied by $\ll$ in* (3.11). *Then we have as $M \to \infty$*

$$P\{|\sum_{j\leqq M} \bar{D}_j^*| \geqq 8RB(M^{\varepsilon+1}\log\log M^{\varepsilon+1})^{1/2}\} \ll \exp(-12RB\log\log M^{\varepsilon+1}).$$

Proof. Put

$$u_n = \sum_{j\leqq n} \bar{D}_j^*, \quad n \leqq M$$
$$= u_M, \qquad n > M;$$
$$s_n^2 = \sum_{j\leqq n} E(\bar{D}_j^{*2}|D_1, \ldots, D_{j-1}), \quad n \leqq M$$
$$= s_M^2, \qquad n > M;$$
$$c = 2t_M, \quad \lambda = 2(\log\log M^{\varepsilon+1})^{1/2}M^{-(\varepsilon+1)/2}, \quad K = 4RBM^{\varepsilon+1},$$
$$T_n = \exp(\lambda u_n - (1/2)\lambda^2(1+(1/2)\lambda c)s_n^2).$$

Thus $\{u_n, n \geqq 1\}$ is a martingale. Moreover, $\bar{D}_j^* \leqq c$ a.s. and $\lambda c \leqq 1$. Now the proof can be finished as in the proof of Lemma 6.

Lemma 18.

$$P\{|\sum_{j\leqq M} E(D_j^*|D_1, \ldots, D_{j-1})| \geqq 2RB(M^{\varepsilon+1}\log\log M^{\varepsilon+1})^{1/2}\} \ll R^{-1}M^{-1}(\log\log M).$$

Proof. The probability in question is, by (3.10),

$$\leqq P\{|\sum_{j\leqq M} E(D_j|D_1, \ldots, D_{j-1})| \geqq RB(M^{\varepsilon+1}\log\log M^{\varepsilon+1})^{1/2}\} +$$
$$+ P\{\sum_{j\leqq M} E(D_j I[|D_j| > t_M]|D_1, \ldots, D_{j-1}) \geqq RB(M^{\varepsilon+1}\log\log M^{\varepsilon+1})^{1/2}\} \leqq$$
$$\leqq 0 + \sum_{j\leqq M} E|E(D_j I[|D_j| > t_M]|D_1, \ldots, D_{j-1})|/RM^{(\varepsilon+1)/2}(\log\log M)^{1/2} \leqq$$
$$\leqq \sum_{j\leqq M} E^{1/4}(D_j^4)P^{3/4}\{|D_j| > t_M\}/RM^{(\varepsilon+1)/2}(\log\log M)^{1/2} \ll$$
$$\ll \sum_{j\leqq M} E^{1/2}(T_j^4)E^{3/4}(D_j^4)t_M^{-3}/RM^{(\varepsilon+1)/2}(\log\log M)^{1/2} \ll$$
$$\ll \sum_{j\leqq M} j^{\varepsilon/2}j^{3\varepsilon/2}M^{-3(\varepsilon+1)/2}(\log\log M)^{3/2}/RM^{(\varepsilon+1)/2}(\log\log M)^{1/2} \ll$$
$$\ll R^{-1}M^{2\varepsilon+1-2(\varepsilon+1)}(\log\log M) = R^{-1}M^{-1}(\log\log M)$$

by Lemma 16(1), Hölder's inequality, Lemma 13, Markov inequality, and Lemma 11.

Lemma 19.

$$P\Big\{\Big|\sum_{j\leqq M}\bar{D}_j\Big|\geqq 12RB(M^{\varepsilon+1}\log\log M^{\varepsilon+1})^{1/2}\Big\}\ll$$

$$\ll \exp(-12RB\log\log M)+R^{-1}M^{-1}(\log\log M)+R^{-2}M^{-1/2}(\log\log M).$$

Proof. Put $\lambda=10RB(M^{\varepsilon+1}\log\log M^{\varepsilon+1})^{1/2}$. Now the probability in question is

$$\leqq P\Big\{\Big|\sum_{j<M}D_j\Big|>\lambda\Big\}=$$

$$=P\Big\{\Big|\sum_{j\leqq M}D_j\Big|>\lambda;\ \max_{j\leqq M}|D_j|\leqq t_M\Big\}+P\Big\{\Big|\sum_{j\leqq M}D_j\Big|>\lambda;\ \max_{j\leqq M}|D_j|>t_M\Big\}\leqq$$

$$\leqq P\Big\{\Big|\sum_{j\leqq M}D_j^*\Big|>\lambda\Big\}+P^{1/2}\Big\{\Big|\sum_{j\leqq M}D_j\Big|>\lambda\Big\}P^{1/2}\Big\{\max_{j\leqq M}|D_j|>t_M\Big\}\leqq$$

$$\leqq P\Big\{\Big|\sum_{j\leqq M}\bar{D}_j^*\Big|>8RB(M^{\varepsilon+1}\log\log M^{\varepsilon+1})^{1/2}\Big\}+$$

$$+P\Big\{\Big|\sum_{j\leqq M}E(D_j^*|D_1,\ldots,D_{j-1})\Big|>2RB(M^{\varepsilon+1}\log\log M^{\varepsilon+1})^{1/2}\Big\}+$$

$$+P^{1/2}\Big\{\Big|\sum_{j\leqq M}T_j\Big|>\frac{\lambda}{2}\Big\}\Big(\sum_{j\leqq M}P\{|D_j|>t_M\}\Big)^{1/2}\ll$$

$$\ll \exp(-12RB\log\log M)+R^{-1}M^{-1}(\log\log M)+$$

$$+(M^{2(\varepsilon+1)}/R^4M^{2(\varepsilon+1)})^{1/2}\Big(t_M^{-4}\sum_{j\leqq M}j^{2\varepsilon}\Big)^{1/2}\ll$$

$$\ll \exp(-12RB\log\log M)+R^{-1}M^{-1}(\log\log M)+R^{-2}M^{-1/2}(\log\log M)$$

by (3.9), (3.10), Lemma 17, Lemma 18, Hölder's inequality, and Markov inequality.

Lemma 20. *There is a positive constant A_1 ($\geqq 1$) such that*

$$P\Big\{\Big|\sum_{j\leqq M}T_j\Big|\geqq A_1R(M^{\varepsilon+1}\log\log M^{\varepsilon+1})^{1/2}\Big\}\ll$$

$$\ll \exp(-12R\log\log M^{\varepsilon+1})+(R^{-1}M^{-1}+R^{-2}M^{-1/2})\log\log M.$$

The proof is very simple (see Lemma 7). Similarly,

$$P\Big\{\Big|\sum_{j\leqq M}T_j'\Big|\geqq A_1R(M^{\varepsilon+1}\log\log M^{\varepsilon+1})^{1/2}\Big\}\ll$$

$$\ll \exp(-12R\log\log M^{\varepsilon+1})+(R^{-1}M^{-1}+R^{-2}M^{-1/2})\log\log M$$

where

$$T_j'=\sum_{v\in I_j}\cos 2\pi n_v x.$$

Lemma 21.

$$P\Big\{\Big|\sum_{k\leqq N}\cos 2\pi n_k x\Big|\geqq 3A_1R(N\log\log N)^{1/2}\Big\}\ll$$

$$\ll \exp(-11R\log\log N)+(R^{-1}N^{-1/(\varepsilon+1)}+R^{-2}N^{-1/2(\varepsilon+1)})\log\log N.$$

Proof. Assume first that $N \in H_{M+1}$ for some M. Then the lemma follows from the following estimation:

$$P\{|\sum_{v=C_{M+1}}^{N} \cos 2\pi n_v x| \geqq AR(N \log\log N)^{1/2}\} \ll M^{2\varepsilon}/R^4 N^2 \ll R^{-4} N^{-2/(1+\varepsilon)}$$

by Markov inequality and Lemma 11.

For the case $N \in I_{M+1}$, the proof is the same as above.

Similar to Lemma 21, one can prove that

$$P\{|\sum_{k \leqq N} \sin 2\pi n_k x| \geqq 3A_1 R(N \log\log N)^{1/2}\} \ll \tag{3.12}$$

$$\ll \exp(-11R \log\log N) + (R^{-1} N^{-1/(\varepsilon+1)} + R^{-2} N^{-1/2(\varepsilon+1)}) \log\log N.$$

It is clear now that Proposition 3 follows from Lemma 21 and (3.12).

3.2 Proof of Proposition 4.

Theorem A (Theorem 12.2 of [5]). *Let $\xi_1, \dots, \xi_m$ be random variables. Let $s_k = \xi_1 + \dots + \xi_k$ ($s_0 = 0$), and put $M_m = \max_{0 \leqq k \leqq m} |s_k|$. Suppose for some $\gamma \geqq 0$, $\alpha > 1$, and some $u_1, \dots, u_m > 0$,*

$$P\{|s_j - s_i| \geqq \lambda\} \leqq (1/\lambda^\gamma)(\sum_{i<\varrho\leqq j} u_\varrho)^\alpha, \quad 0 \leqq i < j \leqq m$$

for all $\lambda > 0$. Then for all $\lambda > 0$,

$$P\{M_m > \lambda\} \leqq (k/\lambda^\gamma)(\sum_{0<\varrho\leqq m} u_\varrho)^\alpha$$

where K is a constant depending only on γ and α.

Lemma 22. *For $k \geqq 1$, put $Z_k = \sum_{v=N+1}^{N+k} \cos 2\pi n_v x$. Then, for some $c_0 \geqq 2$, we have*

$$P\{\max_{1\leqq k\leqq N} |Z_k| \geqq 24 A_1 R C_0 (N \log\log N)^{1/2}\} \ll$$

$$\ll \exp(-11R \log\log N) + (R^{-1} N^{-1/(\varepsilon+1)} + R^{-2} N^{-1/2(\varepsilon+1)}) \log\log N$$

where A_1 is the constant appearing in Lemma 20.

Note that under the condition (1.2) of $\{n_k\}$, some parts of the proof of Lemma 8 (or Lemma (6) of [1]) need to be changed.

Proof of Lemma 22. Suppose that $N+1 \in H_m$ and $2N \in H_M$ for some m and M. (The proofs for other cases are the same.)

Let t be any number $\geqq 3\sqrt{N}$. Thus

$$P\{\max_{1\leqq k\leqq N}|Z_k|>6t\}\leqq P\{\max_{m\leqq k\leqq M}|\sum_{j=m}^{k}T_j|>2t\}+$$

$$+P\{\max_{m\leqq k\leqq M}\max_{j\in H_k}|\sum_{v=c_k}^{j}\cos 2\pi n_v x|>t\}+P\{\max_{m\leqq k\leqq M-1}|\sum_{j=m}^{k}T_j'|>2t\}+$$

$$+P\{\max_{m\leqq k\leqq M-1}\max_{j\in I_k}|\sum_{v=d_k+1}^{j}\cos 2\pi n_v x|>t\}=I_1+I_1'+I_2+I_2' \quad \text{(say)}.$$

By Theorem A, Lemma 11 together with Markov inequality, we have

$$I_1'\leqq\sum_{k=m}^{M}P\{\max_{j\in H_k}|\sum_{v=c_k}^{j}\cos 2\pi n_v x|>t\}\ll$$

$$\ll\sum_{k=m}^{M}k^{2\varepsilon}/t^4\ll M^{2\varepsilon+1}/t^4\ll N^{(2\varepsilon+1)/(\varepsilon+1)}/t^4.$$

Similarly, $I_2'\ll N^{(2\varepsilon+1)/(\varepsilon+1)}/t^4$. By Lemma 16(1),

$$I_1\leqq P\{\max_{m\leqq k\leqq M}|\sum_{j=m}^{k}\bar{D}_j|>t\}+P\{\max_{m\leqq k\leqq M}|\sum_{j=m}^{k}(T_j-\bar{D}_j)|>t\}\leqq \tag{3.13}$$

$$\leqq P\{\max_{m\leqq k\leqq M}|\sum_{j=m}^{k}\bar{D}_j|>t\}.$$

Write $\bar{Z}_k=\sum_{j=m}^{k}\bar{D}_j\ (m\leqq k\leqq M)$, and put

$$A=\{\max_{m\leqq k\leqq M}\bar{Z}_k>t\},\quad A_m=\{\bar{Z}_m>t\},$$

$$A_k=\{\bar{Z}_m\leqq t,\ \bar{Z}_{m+1}\leqq t,\ \dots,\ \bar{Z}_{k-1}\leqq t,\ \bar{Z}_k>t\}\quad (m<k\leqq M),$$

$$B_k=\{\bar{Z}_m-\bar{Z}_k>-C_0\sqrt{N}\}\quad (m\leqq k\leqq M),\quad C=\{\bar{Z}_M>t-C_0\sqrt{N}\}.$$

Thus A_kB_k are pairwise disjoint. Also $\bigcup_{k=m}^{M}A_kB_k\subset C$.

On B_k^c we have $(\bar{Z}_M-\bar{Z}_k)^2>C_0^2N$. Thus

$$P(A_kB_k^c)=\int_{A_k}1_{B_k^c}\leqq C_0^{-2}N^{-1}\int_{A_k}(\bar{Z}_M-\bar{Z}_k)^2= \tag{3.14}$$

$$=C_0^{-2}N^{-1}\sum_{\substack{I\subset A_k\cap\sigma(D_1,\dots,D_k)\\ P(I)=2^{-md_k}}}\int_I(\bar{Z}_M-\bar{Z}_k)^2.$$

For each such I in the summation,

(3.15)

$$\int_I (\bar{Z}_M-\bar{Z}_k)^2 = \int_I E((\bar{Z}_M-\bar{Z}_k)^2|D_1, \dots, D_k) = \int_I E((\sum_{j=k+1}^{M} \bar{D}_j)^2|D_1, \dots, D_k) \leqq$$

$$\leqq \|E((\sum_{j=k+1}^{M} \bar{D}_j)^2|D_1, \dots, D_k)\|_\infty P(I) = \|E(\sum_{j=k+1}^{M} \bar{D}_j^2|D_1, \dots, D_k)\|_\infty P(I) \leqq$$

$$\leqq (\sum_{j=k+1}^{M} j^s)P(I) \leqq C_0 NP(I) \quad \text{for some } C_0.$$

Hence (3.14) and (3.15) imply

$$P(A_k B_k^c) \leqq C_0^{-2} N^{-1}(C_0 N)\sum_I P(I) = C_0^{-1}P(A_k) \leqq (1/2)P(A_k).$$

Now

$$(1/2)P(A) = (1/2)\sum_{k=m}^{M} P(A_k) \leqq \sum_{k=m}^{M} (P(A_k)-P(A_k B_k^c)) = \sum_{k=m}^{M} P(A_k B_k) \leqq P(C).$$

This proves that

$$P\{\max_{m\leqq k\leqq M} \sum_{j=m}^{M} \bar{D}_j > t\} \leqq 2P\{\sum_{j=m}^{M} \bar{D}_j > t-C_0\sqrt{N}\}.$$

Similarly,

$$P\{\max_{m\leqq k\leqq M} -\sum_{j=m}^{k} \bar{D}_j > t\} \leqq 2P\{-\sum_{j=m}^{M} \bar{D}_j > t-C_0\sqrt{N}\}.$$

Hence

$$P\{\max_{m\leqq k\leqq M} |\sum_{j=m}^{k} \bar{D}_j| > t\} \leqq 2P\{|\sum_{j=m}^{M} \bar{D}_j| > t-C_0\sqrt{N}\}.$$

Therefore, by (3.13)

$$I_1 \leqq 2P\{|\sum_{j=m}^{M} \bar{D}_j| > t-C_0\sqrt{N}\}.$$

But

$$P\{|\sum_{j=m}^{M} \bar{D}_j| > t-C_0\sqrt{N}\} \leqq P\{|\sum_{j=m}^{M} T_j| > t/2-(C_0/2)\sqrt{N}\}.$$

Hence

$$I_1 \ll P\{|\sum_{j=m}^{M} T_j| > t/2-(C_0/2)\sqrt{N}\}.$$

Similarly,

$$I_2 \ll P\{|\sum_{j=m}^{M-1} T_j| > t/2-(C_0/2)\sqrt{N}\}.$$

We can conclude now that if $t=4ARC_0(N\log\log N)^{1/2}$, then

$$P\Big\{\max_{m\le k\le M}|Z_k|>6t\Big\}\ll$$

$$\ll P\Big\{\Big|\sum_{j=m}^{M}T_j\Big|>t/2-(C_0/2)\sqrt{N}\Big\}+P\Big\{\Big|\sum_{j=m}^{M-1}T_j'\Big|>t/2-(C_0/2)\sqrt{N}\Big\}+N^{(2\varepsilon+1)/(\varepsilon+1)}/t^4\ll$$

$$\ll P\Big\{\Big|\sum_{j=1}^{M}T_j\Big|>t/4-(C_0/4)\sqrt{N}\Big\}+P\Big\{\Big|\sum_{j=1}^{m-1}T_j\Big|>t/4-(C_0/4)\sqrt{N}\Big\}+$$

$$+P\Big\{\Big|\sum_{j=1}^{M-1}T_j'\Big|>t/4-(C_0/4)\sqrt{N}\Big\}+P\Big\{\Big|\sum_{j=1}^{m-1}T_j'\Big|>t/4-(C_0/4)\sqrt{N}\Big\}+N^{(2\varepsilon+1)/(\varepsilon+1)}/t^4\ll$$

$$\ll\exp(-11R\log\log N)+(R^{-1}N^{1/(\varepsilon+1)}+R^{-2}N^{-1/2(\varepsilon+1)})\log\log N+N^{(2\varepsilon+1)/(\varepsilon+1)}t^{-4}\ll$$

$$\ll\exp(-11R\log\log N)+(R^{-1}N^{1/(\varepsilon+1)}+R^{-2}N^{-1/2(\varepsilon+1)})\log\log N.$$

Putting $A>24A_1C_0$, Proposition 4 is proved. We choose A large enough that both Proposition 3 and 4 will apply.

4. Further results

Recently, I found an improvement of Theorem 2. The proof requires only little work beyond the one of Theorem 2. We shall give a sketch of the proof.

Theorem 3. *Suppose that the sequence $\{n_k\}$ of integers satisfies (1.2) and furthermore, for any $v>0$, the number of solutions of $n_k\pm n_l=v$ $(1\le k, l\le N)$ is at most BN^η with constants $B>0$, $\eta<(1-\delta)/2$. Then for each α with $\alpha>1/2+\delta/2+\eta$,*

$$\limsup_{N\to\infty}\sup_{f\in A_\alpha}\Big|\sum_{k=1}^{N}f(n_kx)\Big|/(N\log\log N)^{1/2}\ll 1$$

for almost all $x\in[0,1]$, where the constant implied by $\ll$ depends on δ, c, B and α.

The case $\eta=0$ in the above theorem means that the sequence $\{n_k\}$ satisfies condition B_2; in this special case we have Theorem 2.

4.1 Proof of Theorem 3. We assume that $\alpha\le 1$. We choose ε so that

(4.1) $$\varepsilon>\delta/(1-\delta)\quad\text{and}\quad 1/2+\varepsilon/2(\varepsilon+1)+\eta<\alpha.$$

(This can be done by choosing ε sufficiently close to $\delta/(1-\delta)$ since $1/2+\delta/2+\eta<\alpha$ and $\delta<1$.) Put

(4.2) $$\gamma=2/(2\alpha+1).$$

Thus, by (4.1),

$$\varepsilon/2(\varepsilon+1) < (1-\gamma)/\gamma \leqq 1/2. \tag{4.3}$$

Now choose ε' so that

$$\max\left(\frac{\varepsilon}{2(\varepsilon+1)}, \frac{2\varepsilon+1}{8(\varepsilon+1)}+\frac{\eta}{2}, \frac{1}{2}-\frac{1}{2(\varepsilon+1)}+\eta\right) < \varepsilon' < \tag{4.4}$$

$$< \min\left(\frac{1-\gamma}{\gamma}, 1-\frac{\delta}{2}-\frac{1}{2(\varepsilon+1)}\right).$$

(It is easy to verify that max<min in (4.4).) We choose β so that

$$1/2\varepsilon' > \beta > \gamma/2(1-\gamma) \tag{4.5}$$

and put

$$q = \beta/(\beta-1). \tag{4.6}$$

Thus, by (4.4) and (4.5), we have

$$-q+2q\varepsilon' < -1. \tag{4.7}$$

We also have

$$1-2\varepsilon' < (1-2\eta(\varepsilon+1))/(\varepsilon+1) \tag{4.8}$$

and

$$1-4\varepsilon' < 1-(2\varepsilon+1)/2(\varepsilon+1)-2\eta$$

by (4.4).

We define the events, for integers h and $n \geqq 0$,

$$G(n, h) = \{F(0, 2^n, h) \geqq A|h|^{2\varepsilon'}\varphi(2^n)\};$$

$$G_n = \bigcup_{1\leqq|h|\leqq 2^n} G(n, h);$$

$$H(n, h) = \{\bigcup_{1\leqq m<2^n} F(2^n, m, h) \geqq A|h|^{2\varepsilon'}\varphi(2^n)\};$$

$$H_n = \bigcup_{1\leqq|h|\leqq 2^n} H(n, h)$$

where A is the constant appearing in Proposition 5 below.

Proposition 5. *Let $\{n_k\}$ be as in Theorem 3. Then there exist positive constants A and C such that*

$$P\{|\sum_{k=1}^{N} \exp(2\pi i n_k x)| \geqq AR(N \log\log N)^{1/2}\} \leqq$$

$$\leqq C[\exp(-10R\log\log N)+R^{-1}N^{-(1-2\eta(\varepsilon+1))/(\varepsilon+1)}\log\log N+$$

$$+R^{-2}N^{-(1-(2\varepsilon+1)/2(\varepsilon+1)-2\eta)}\log\log N]$$

for all $R\geqq 1$, $N\geqq 1$.

Proposition 6. *Let $\{n_k\}$ be as in Theorem 3. Then*

$$P\{\max_{1\leqq m<N}\Big|\sum_{k=N+1}^{N+m}\exp(2\pi i n_k x)\Big|\geqq AR(N\log\log N)^{1/2}\}\leqq$$
$$\leqq C\,[\exp(-10R\log\log N)+R^{-1}N^{-(1-2\eta(\varepsilon+1))/(\varepsilon+1)}\log\log N+$$
$$+R^{-2}N^{-(1-(2\varepsilon+1)/2(\varepsilon+1)-2\eta)}\log\log N]$$

for all $R\geqq 1$, $N\geqq 1$.

Here A and C are the same constants as in Proposition 5.

Taking Proposition 5 and 6 for granted we see that

$$P(G_n)\ll\sum_{1\leqq|h|\leqq 2^n}[\exp(-2|h|^{2\varepsilon'}\log n)+$$
$$+|h|^{-2\varepsilon'}2^{-(1-2\eta(\varepsilon+1))n/(\varepsilon+1)}\log n+|h|^{-4\varepsilon'}2^{-(1-(2\varepsilon+1)/2(\varepsilon+1)-2\eta)n}\log n]\ll$$
$$\ll n^{-2}+2^{(1-2\varepsilon'-(1-2\eta(\varepsilon+1))/(\varepsilon+1))n}\log n+2^{[1-4\varepsilon'-(1-(2\varepsilon+1)/2(\varepsilon+1)-2\eta)]n}\log n;$$
$$P(H_n)\ll n^{-2}+2^{(1-2\varepsilon'-(1-2\eta(\varepsilon+1))/(\varepsilon+1))n}\log n+2^{[1-4\varepsilon'-(1-(2\varepsilon+1)/2(\varepsilon+1)-2\eta)]n}\log n.$$

Since $1-2\varepsilon'<(1-2\eta(\varepsilon+1))/(\varepsilon+1)$ and $1-4\varepsilon'<1-(2\varepsilon+1)/2(\varepsilon+1)-2\eta$ by (4.8), we obtain

$$\sum_{n\geqq 0}P(G_n)<\infty,\quad \sum_{h\geqq 0}P(H_n)<\infty.$$

Thus, by Borel—Cantelli Lemma,

$$F(0,N,h)\ll|h|^{2\varepsilon'}\varphi(N)\quad\text{a.s.}$$

for all $1\leqq|h|\leqq N/2$. Now (3.7) can be obtained using ε', β and q in (4.4), (4.5) and (4.6) respectively.

Since the proofs of Proposition 5 and 6 are similar to those of Proposition 3 and 4 of Section 3, we omit them.

Acknowledgement. I am deeply grateful to Professor Walter Philipp, my advisor, for his guidance, assistance and encouragement throughout the preparation of this paper. I also wish to thank Professor Robert Kaufman for his several useful suggestions which contributed to the development of the paper.

References

[1] I. Berkes, An almost sure invariance principle for lacunary trigonometric series, *Acta Math. Hungar.*, **26** (1975), 209—220.

[2] I. Berkes, On the asymptotic behaviour of $\Sigma f(n_k x)$. I—II, *Z. Wahrsch. Verw. Gebiete*, **41** (1977), 115—137.

[3] I. Berkes, On the central limit theorems of lacunary trigonometric series, *Anal. Math.*, **4** (1978), 159—180.

[4] I. Berkes, A central limit theorem for trigonometric series with small gaps, *Z. Wahrsch. Verw. Gebiete*, **47** (1979), 157—161.

[5] P. Billingsley, *Convergence of probability measures*, Wiley (New York, 1968).

[6] S. Dhompongsa, Limit theorems for weakly dependent random vectors, Ph. D. Thesis, University of Illinois, Urbana, 1982.

[7] P. Erdős, On trigonometric sums with gaps, *Magyar Tud. Akad. Mat. Kut. Int. Közl.*, **7A** (1962), 37—42.

[8] V. G. Gaposhkin, Lacunary series and independent functions, *Russian Math. Surveys*, **22** (1966), 3—82.

[9] M. Kac, R. Salem and A. Zygmund, A gap theorem, *Trans. Amer. Math. Soc.*, **63** (1948), 235—243.

[10] R. Kaufman and W. Philipp, A uniform law of the iterated logarithm for classes of functions, *Ann. Probab.*, **6**(6) (1978), 930—952.

[11] W. Philipp, A functional law of the iterated logarithm for empirical distribution functions of weakly dependent random variables, *Ann. Probab.*, **5**(3) (1977), 319—350.

[12] W. F. Stout, *Almost sure convergence*, Academic Press (New York, 1974).

[13] S. Takahachi, An asymptotic property of a gap sequence, *Proc. Japan Acad.*, **38** (1962), 101—104.

[14] S. Takahachi, The law of the iterated logarithm for a gap sequence with infinite gaps, *Tôhoku Math. J.*, **15** (1963), 281—288.

[15] A. Zygmund, *Trigonometric series* (Warszawa—Lwów, 1935).

[16] A. Zygmund, *Trigonometric series*, Vol. II. Cambridge University Press (1959).

UNIVERSITY OF ILLINOIS
URBANA, IL. 61801, U.S.A.
and
CHIANG MAI UNIVERSITY
CHIANG MAI, THAILAND

Acta Sci. Math., 50 (1986), 125—131

Toeplitz-Kriterien für Klassen von Matrixabbildungen zwischen Räumen stark limitierbarer Folgen

EBERHARD MALKOWSKY

1. Allgemeine Sätze

In [1] haben wir in Verbindung mit stark limitierbaren Folgen als Verallgemeinerung der Räume w_p $(0<p<\infty)$ von MADDOX [2] die Räume $[\tilde{C}]$ für alle $\alpha>0$ definiert, wobei $[\tilde{C}]$ mit den Bezeichnungen von [1] für die folgenden Mengen geschrieben worden ist: $(0<p<\infty)$

$$[\tilde{C}_\alpha]^p := \left\{x\in s \mid \text{Es gibt ein } l\in \mathbf{C} \text{ mit } \left(\|(x-le)^{(\nu)}(\alpha;\, p)\|_p\right)_\nu \in c_0\right\},$$

$$[\tilde{C}_\alpha]_0^p := \left\{x\in s \,\middle|\, \left(\|x^{(\nu)}(\alpha;\, p)\|_p\right)_\nu \in c_0\right\} \quad \text{und} \quad [\tilde{C}_\alpha]_\infty^p := \left\{x\in s \,\middle|\, \left(\|x^{(\nu)}(\alpha;\, p)\|_p\right)_\nu \in l_\infty\right\},$$

versehen mit der p-Norm für $0<p<1$ bzw. der Norm für $1\leqq p<\infty$

$$\|x\|_{[\tilde{C}]} := \left\|\left(\|x^{(\nu)}(\alpha;\, p)\|_p\right)_\nu\right\|_\infty \quad \text{für alle} \quad x\in[\tilde{C}]$$

(s. [1], Definition 2.1, Definition 2.2).

Weiter haben wir in einigen Fällen Toeplitz-Kriterien für Klassen (X, Y) unendlicher komplexer Matrizen erhalten, die den Folgenraum X in den Folgenraum Y abbilden, wobei entweder $X=[\tilde{C}]$ oder $Y=[\tilde{C}]$.

Hier wollen wir Toeplitz-Kriterien für (X, Y) in dem Fall herleiten, in dem $X=[\tilde{C}]$ und $Y=[\tilde{C}]$. Wir benutzen dabei immer die in [1] eingeführten Bezeichnungen. Es seien stets $\alpha, \beta>0$.

Wir beweisen zunächst drei allgemeine Resultate, aus denen unter anderem die Toeplitz-Kriterien für die Klasse $([\tilde{C}_\alpha]^p, [\tilde{C}_\beta]^1)$ folgen.

Satz 1.1. *Sei X ein vollständiger r-normierter Teilraum von s. Gilt $X^\dagger\subset X^*$, so folgt: $A\in(X, [\tilde{C}_\alpha]_\infty^1)\Leftrightarrow M(\dagger, \alpha)<\infty$, wobei*

$$M(\dagger, \alpha) := \sup_{\mu\in \mathbf{N}_0} \left\{\max_{N_\mu\subset N^{(\mu)}} \left\|\left((1/A_{2^\mu-1}^\alpha) \sum_{n\in N_\mu} A_{2^{\mu+1}-n}^{\alpha-1} a_{nk}\right)_k\right\|^\dagger\right\} < \infty \quad (\alpha > 0).$$

Eingegangen am 10. November 1983.

Beweis. Es gelte $A \in (X, [\tilde{C}_\alpha]^1_\infty)$. Dann existiert $A_n(x) := \sum_{k=1}^{\infty} a_{nk} x_k$ für alle $x \in X$ und für alle $n \in \mathbf{N}$, d. h. $(a_{nk})_k \in X^\dagger$ für alle $n \in \mathbf{N}$, und wegen $X^\dagger \subset X^*$ folgt $A_n \in X^*$ für alle $n \in \mathbf{N}$. Für alle $\mu \in \mathbf{N}_0$ und für alle $N_\mu \subset N^{(\mu)}$ setzen wir

$$B^\alpha_{N_\mu} := (1/A^\alpha_{2^\mu - 1}) \sum_{n \in N_\mu} A^{\alpha-1}_{2^{\mu+1} - n} A_n \quad \text{und} \quad b^\alpha_{N_\mu, k} := (1/A^\alpha_{2^\mu - 1}) \sum_{n \in N_\mu} A^{\alpha-1}_{2^{\mu+1} - n} a_{nk}$$
$$(k = 1, 2, \ldots).$$

Sei $\mu \in \mathbf{N}_0$ beliebig. Dann gilt für alle $N_\mu \subset N^{(\mu)}$ und für alle $x \in S_X$ wegen $B^\alpha_{N_\mu} \in X^*$

$$|B^\alpha_{N_\mu}(x)| = \left|\sum_{k=1}^{\infty} b^\alpha_{N_\mu, k} x_k\right| \leqq \|B^\alpha_{N_\mu}\|$$

und daher $\|(b^\alpha_{N_\mu, k})_k\|^\dagger \leqq \|B^\alpha_{N_\mu}\|$ für alle $N_\mu \subset N^{(\mu)}$, also für alle $\mu \in \mathbf{N}_0$:

$$\|(b^\alpha_{N'_\mu, k})_k\|^\dagger := \max_{N_\mu \subset N^{(\mu)}} \|(b^\alpha_{N_\mu, k})_k\|^\dagger \leqq \|B^\alpha_{N''_\mu}\| := \max_{N_\mu \subset N^{(\mu)}} \|B^\alpha_{N_\mu}\|.$$

Wegen $A \in (X, [\tilde{C}_\alpha]^1_\infty)$ gilt für alle $x \in X$

$$\varlimsup_{\mu \to \infty} |B^\alpha_{N''_\mu}(x)| \leqq \varlimsup_{\mu \to \infty} \left\|((A_n(x))^{(\mu)}(\alpha;\, 1))_n\right\|_1 \leqq \left\|(A_n(x))_n\right\|_{[C]} < \infty.$$

Mit dem Satz von Banach—Steinhaus (s. [3], Satz 12, S. 115, bzw. [2], S. 286) folgt

$$\sup_{\mu \in \mathbf{N}_0} \|(b^\alpha_{N'_\mu, k})_k\|^\dagger = M(\dagger, \alpha) \leqq \sup_{\mu \in \mathbf{N}_0} \|B^\alpha_{N''_\mu}\| < \infty.$$

Umgekehrt gelte $M(\dagger, \alpha) < \infty$.

Seien $n \in \mathbf{N}$ und $x \in X$ beliebig. Dann gibt es ein $\mu_n \in \mathbf{N}_0$ mit $n \in N^{(\mu_n)}$. Wähle $N_{\mu_n} := \{n\}$. Wegen $M(\dagger, \alpha) < \infty$ und der Definition von $\|\cdot\|^\dagger$ existiert dann

$$(1/A^\alpha_{2^{\mu_n} - 1}) A^{\alpha-1}_{2^{\mu_n}+1-n} A_n(x)$$

und damit auch $A_n(x)$ für alle $n \in \mathbf{N}$ und für alle $x \in X$. Sei $x \in X$ beliebig. Dann gilt für alle $\mu \in \mathbf{N}_0$ und für alle $N_\mu \subset N^{(\mu)}$

$$|B^\alpha_{N_\mu}(x)| \leqq \|(b^\alpha_{N_\mu, k})\|^\dagger \|x\|^{1/r} \leqq \max_{N_\mu \subset N^{(\mu)}} \|(b^\alpha_{N_\mu, k})_k\|^\dagger \|x\|^{1/r} =: M(\mu) \|x\|^{1/r}$$

und damit für alle $\mu \in \mathbf{N}_0$ mit einer bekannten Ungleichung (s. [4], S. 33)

$$\left\|((A_n(x))^{(\mu)}(\alpha;\, 1))_n\right\|_1 \leqq 4 \cdot \max_{N_\mu \subset N^{(\mu)}} |B^\alpha_{N_\mu}(x)| \leqq 4 \cdot M(\mu) \|x\|^{1/r},$$

d. h.

$$\text{(1.1)} \qquad \left\|(A_n(x))_n\right\|_{[C]} = \left\|(\|((A_n(x))^{(\mu)}(\alpha;\, 1))_n\|_1)_\mu\right\|_\infty \leqq$$
$$\leqq \sup_{\mu \in \mathbf{N}_0} 4 \cdot M(\mu) \|x\|^{1/r} = 4 \cdot M(\dagger, \alpha) \cdot \|x\|^{1/r} < \infty.$$

Damit ist der Satz bewiesen.

Satz 1.2. *Sei X ein vollständiger r-normierter Teilraum von s mit Basis $(e^{(k)})_k$. Dann gilt für alle* $\alpha>0$

$$A\in(X,[\tilde{C}_\alpha]^1)\Leftrightarrow\begin{cases}\text{(i)} & M(\dagger,\alpha)<\infty \quad (M(\dagger,\alpha) \textit{ wie in Satz } 1.1)\\ \text{(ii)} & (a_{nk})_n\in[\tilde{C}_\alpha]^1 \textit{ für alle } k=1,2,\dots .\end{cases}$$

Beweis. Wie im Beweis von Satz 1.1 in [1] folgt $X^\dagger\subset X^*$. Es gelte $A\in(X,[\tilde{C}_\alpha]^1)$. Wegen $[\tilde{C}_\alpha]^1\subset[\tilde{C}_\alpha]^1_\infty$ folgt $A\in(X,[\tilde{C}_\alpha]^1_\infty)$ und daher (i) mit Satz 1.1. Weiter ist $(A_n(x))_n\in[\tilde{C}_\alpha]^1$ für alle $x\in X$, d. h. insbesondere gilt für jedes $x=e^{(k)}\in X$ $(k=1,2,\dots)$ $(A_n(e^{(k)}))_n=(a_{nk})_n\in[\tilde{C}_\alpha]^1$.

Umgekehrt seien (i) und (ii) erfüllt. Aus (i) folgt wie im Beweis von Satz 1.1 die Existenz von $A_n(x):=\sum_{k=1}^\infty a_{nk}x_k$ für alle $x\in X$ und für alle $n\in\mathbf{N}$. Wegen (ii) gibt es zu jedem $k=1,2,\dots$ ein $a_k\in\mathbf{C}$ mit

$$\lim_{\mu\to\infty}\|((a_{nk}-a_k)^{(\mu)}(\alpha;\ 1))_n\|_1=0. \tag{1.2}$$

Es gilt dann $(a_k)_k\in X^\dagger$. Denn: Sei $x=\sum_{k=1}^\infty x_k e^{(k)}\in X$ beliebig; sei $\varepsilon>0$ gegeben. Wähle $k_0\in\mathbf{N}$ so groß, daß für alle l, $m>k_0$ $(m>l)$ mit

$$x^{(l,m)}:=\sum_{k=l}^m x_k e^{(k)} \quad \text{gilt} \quad \|x^{(l,m)}\|<(\varepsilon/(4\cdot M(\dagger,\alpha)+1))^r.$$

Wähle nun $\mu_0\in\mathbf{N}_0$, so daß für alle $\mu\geqq\mu_0$ gilt $1-1/A^\alpha_{2^\mu-1}>0$. Dann gilt für alle l, $m>k_0$ $(m>l)$ und für alle $\mu>\mu_0$ mit $S_\mu:=\sum_{k=1}^m(\|((a_k-a_{nk})^{(\mu)}(\alpha,1))_n\|_1|x_k|)$ und der Ungleichung aus [4], S. 33

$$\begin{aligned}(1-(1/A^\alpha_{2^\mu-1}))\Big|\sum_{k=l}^m a_kx_k\Big| &= (1/A^\alpha_{2^\mu-1})\sum_{n=2^\mu+1}^{2^{\mu+1}-1}A^{\alpha-1}_{2^{\mu+1}-n}\Big|\sum_{k=l}^m a_kx_k\Big|\leqq\\ &\leqq\Big\|\Big(\Big(\sum_{k=l}^m a_kx_k\Big)^{(\mu)}(\alpha;\ 1)\Big)_n\Big\|_1\leqq\sum_{k=l}^m\|((a_k-a_{nk})^{(\mu)}(\alpha;\ 1))_n\|_1|x_k|+\\ &\quad+4\cdot\max_{N_\mu\subset N^{(\mu)}}\Big|(1/A^\alpha_{2^\mu-1})\sum_{n\in N_\mu}A^{\alpha-1}_{2^{\mu+1}-n}\sum_{k=l}^m a_kx_k\Big|\leqq\\ &\leqq S_\mu+4\cdot\max_{N_\mu\subset N^{(\mu)}}\Big\|\Big((1/A^\alpha_{2^\mu-1})\sum_{n\in N_\mu}A^{\alpha-1}_{2^{\mu+1}-n}a_k\Big)_k\Big\|^\dagger\|x^{(l,m)}\|^{1/r}<S_\mu+\varepsilon.\end{aligned}$$

Mit (1.2) erhalten wir $\lim_{\mu\to\infty}S_\mu=0$, d. h. es gilt für alle l, $m>k_0$ $(m>l)$

$$\Big|\sum_{k=l}^m a_kx_k\Big|=\lim_{\mu\to\infty}(1-(1/A^\alpha_{2^\mu-1}))\Big|\sum_{k=l}^m a_kx_k\Big|\leqq\varepsilon.$$

Also konvergiert $\sum_{k=1}^\infty a_kx_k$ für alle $x\in X$, d. h. $(a_k)_k\in X^\dagger$. Wegen $X^\dagger\subset X^*$ gilt

$f_a \in X^*$ mit $f_a(x) := \sum_{k=1}^{\infty} a_k x_k$. Sei $x = \sum_{k=1}^{\infty} x_k e^{(k)} \in X$ beliebig. Sei $\varepsilon > 0$ gegeben. Wähle $k_0 \in \mathbf{N}$ so groß, daß für $r_0(x) := \sum_{k=k_0+1}^{\infty} x_k e^{(k)}$ gilt

$$\|r_0(x)\|^{1/r} < \frac{\varepsilon}{2(4 \cdot M(\dagger, \alpha) + 2\|f_a\|) + 1}.$$

Wähle $\mu_0 \in \mathbf{N}_0$, so daß für alle $\mu > \mu_0$

$$\Big\|\Big(\Big(\sum_{k=1}^{k_0} (a_{nk} - a_k) x_k\Big)^{(\mu)}(\alpha;\ 1)\Big)_n\Big\|_1 < \varepsilon/2 \quad \text{und} \quad 1/A^{\alpha}_{2^\mu - 1} < 1.$$

Dann gilt für alle $\mu > \mu_0$ mit der Ungleichung aus [4], S. 33:

$$\Big\|\big((A_n(x) - f_a(x))^{(\mu)}(\alpha;\ 1)\big)_n\Big\|_1 \leqq \Big\|\Big(\Big(\sum_{k=1}^{k_0} (a_{nk} - a_k) x_k\Big)^{(\mu)}(\alpha;\ 1)\Big)_n\Big\|_1 +$$
$$+ \big\|\big((A_n(r_0(x)))^{(\mu)}(\alpha;\ 1)\big)_n\big\|_1 + \big\|\big((f_a(r_0(x)))^{(\mu)}(\alpha;\ 1)\big)_n\big\|_1 <$$
$$< \varepsilon/2 + (4 \cdot M(\dagger, \alpha) + 2 \cdot \|f_a\|)\,\|r_0(x)\|^{1/r} < \varepsilon,$$

d. h. $(A_n(x))_n \in [\tilde{C}_\alpha]^1$ für alle $x \in X$. Damit ist der Satz bewiesen.

Man erhält sofort:

Korollar 1.1. *Sei X ein vollständiger r-normierter Teilraum von s mit Basis $(e^{(k)})_k$. Dann gilt für alle $\alpha > 0$*

$$A \in (X, [\tilde{C}_\alpha]^1_0) \Leftrightarrow \begin{cases} \text{(i)} & M(\dagger, \alpha) < \infty \quad (M(\dagger, \alpha) \textit{ wie in Satz } 1.1) \\ \text{(ii)} & (a_{nk})_n \in [\tilde{C}_\alpha]^1_0 \quad \textit{für alle } k = 1, 2, \ldots. \end{cases}$$

2. Toeplitz-Kriterien für Matrizenklassen bei den Räumen $[\tilde{C}]$

Wir wenden nun die Ergebnisse aus Kapitel 1 an. Zur Vereinfachung der Bezeichnungen setzen wir:

$$M(p;\ \alpha) := \sup_{\mu \in \mathbf{N}_0} \Big\{\max_{N_\mu \subset N^{(\mu)}} \Big\|\Big((1/A^{\alpha}_{2^\mu - 1}) \sum_{n \in N_\mu} A^{\alpha-1}_{2^{\mu+1} - n} a_{nk}\Big)_k\Big\|_q\Big\} \quad \text{für} \quad 0 < p \leqq \infty,$$

wobei

$$\text{(2.1)} \qquad q := \begin{cases} \infty & \text{für} \quad 0 < p \leqq 1 \\ p/(p-1) & \text{für} \quad 1 < p < \infty \\ 1 & \text{für} \quad p = \infty; \end{cases}$$

$$M(\alpha, p;\ \beta) := \sup_{\mu \in \mathbf{N}_0} \Big\{\max_{N_\mu \subset N^{(\mu)}} \Big\|\Big((1/A^{\beta}_{2^\mu - 1}) \sum_{n \in N_\mu} A^{\beta-1}_{2^{\mu+1} - n} a_{nk}\Big)_k\Big\|_{([\tilde{C}_\alpha]^p_\infty)^\dagger}\Big\}$$

für $0 < p < \infty$.

Wir erhalten Sätze, die für alle $\alpha>0$, $\beta>0$ Toeplitz-Kriterien zwischen folgenden Räumen angeben:

(a) aus Räumen absolut limitierbarer Folgen in Räume stark beschränkter bzw. stark limitierbarer Folgen:

Satz 2.1. $A\in(l_p, [\tilde{C}_\alpha]^1_\infty)\Leftrightarrow M(p;\alpha)<\infty$ *für* $0<p\leqq\infty$.

Und im weiteren stets für alle p mit $0<p<\infty$:

Satz 2.2. $A\in(l_p, [\tilde{C}_\alpha]^1_0)\Leftrightarrow\begin{cases}\text{(i)} & M(p;\alpha)<\infty\\ \text{(ii)} & (a_{nk})_n\in[\tilde{C}_\alpha]^1_0 \quad \textit{für alle} \quad k=1,2,\ldots.\end{cases}$

Satz 2.3. $A\in(l_p, [\tilde{C}_\alpha]^1)\Leftrightarrow\begin{cases}\text{(i)} & M(p;\alpha)<\infty\\ \text{(ii)} & (a_{nk})_n\in[\tilde{C}_\alpha]^1 \quad \textit{für alle} \quad k=1,2,\ldots.\end{cases}$

(b) aus Räumen stark beschränkter Folgen in Räume stark beschränkter Folgen:

Satz 2.4. $A\in([\tilde{C}_\alpha]^p_\infty, [\tilde{C}_\beta]^1_\infty)\Leftrightarrow M(\alpha,p;\beta)<\infty$.

(c) aus Räumen stark limitierbarer Folgen in Räume stark beschränkter bzw. stark limitierbarer Folgen:

Satz 2.5. $([\tilde{C}_\alpha]^p_0, [\tilde{C}_\beta]^1_\infty)=([\tilde{C}_\alpha]^p, [\tilde{C}_\beta]^1_\infty)=([\tilde{C}_\alpha]^p_\infty, [\tilde{C}_\beta]^1_\infty)$.

Satz 2.6. $A\in([\tilde{C}_\alpha]^p_0, [\tilde{C}_\beta]^1_0)\Leftrightarrow\begin{cases}\text{(i)} & M(\alpha,p;\beta)<\infty\\ \text{(ii)} & (a_{nk})_n\in[\tilde{C}_\beta]^1_0 \quad \textit{für alle} \quad k=1,2,\ldots.\end{cases}$

Satz 2.7. $A\in([\tilde{C}_\alpha]^p_0, [\tilde{C}_\beta]^1)\Leftrightarrow\begin{cases}\text{(i)} & M(\alpha,p;\beta)<\infty\\ \text{(ii)} & (a_{nk})_n\in[\tilde{C}_\beta]^1 \quad \textit{für alle} \quad k=1,2,\ldots.\end{cases}$

Satz 2.8. $A\in([\tilde{C}_\alpha]^p, [\tilde{C}_\beta]^1_0)\Leftrightarrow\begin{cases}\text{(i)} & M(\alpha,p;\beta)<\infty\\ \text{(ii)} & (a_{nk})_n\in[\tilde{C}_\beta]^1_0 \quad \textit{für alle} \quad k=1,2,\ldots\\ \text{(iii)} & \left(\sum_{k=1}^{\infty} a_{nk}\right)_n\in[\tilde{C}_\beta]^1_0.\end{cases}$

Satz 2.9. $A\in([\tilde{C}_\alpha]^p, [\tilde{C}_\beta]^1)\Leftrightarrow\begin{cases}\text{(i)} & M(\alpha,p;\beta)<\infty\\ \text{(ii)} & (a_{nk})_n\in[\tilde{C}_\beta]^1 \quad \textit{für alle} \quad k=1,2,\ldots\\ \text{(iii)} & \left(\sum_{k=1}^{\infty} a_{nk}\right)_n\in[\tilde{C}_\beta]^1.\end{cases}$

(d) aus Räumen gewöhnlich limitierbarer Folgen in Räume stark limitierbarer Folgen:

Satz 2.10. $A\in(c_0, [\tilde{C}_\alpha]^1_0)\Leftrightarrow\begin{cases}\text{(i)} & M(1;\alpha)<\infty\\ \text{(ii)} & (a_{nk})_n\in[\tilde{C}_\alpha]^1_0 \quad \textit{für alle} \quad k=1,2,\ldots\end{cases}$

Satz 2.11. $A\in(c_0, [\tilde{C}_\alpha]^1)\Leftrightarrow\begin{cases}\text{(i)} & M(1;\alpha)<\infty\\ \text{(ii)} & (a_{nk})_n\in[\tilde{C}_\alpha]^1 \quad \textit{für alle} \quad k=1,2,\ldots.\end{cases}$

Satz 2.12. $A\in(c,[\tilde{C}_\alpha]_0^1) \Leftrightarrow \begin{cases} \text{(i)} & M(1;\,\alpha)<\infty \\ \text{(ii)} & (a_{nk})_n\in[\tilde{C}_\alpha]_0^1 \quad \textit{für alle} \quad k=1,2,\dots \\ \text{(iii)} & \left(\sum_{k=1}^{\infty} a_{nk}\right)_n\in[\tilde{C}_\alpha]_0^1. \end{cases}$

Satz 2.13. $A\in(c,[\tilde{C}_\alpha]^1) \Leftrightarrow \begin{cases} \text{(i)} & M(1;\,\alpha)<\infty \\ \text{(ii)} & (a_{nk})_n\in[\tilde{C}_\alpha]^1 \quad \textit{für alle} \quad k=1,2,\dots \\ \text{(iii)} & \left(\sum_{k=1}^{\infty} a_{nk}\right)_n\in[\tilde{C}_\alpha]^1. \end{cases}$

Bemerkung. Satz 2.9 enthält im Fall $\alpha=\beta=1$ einen bekannten Satz von KUTTNER und THORPE [5] (Satz 1).

Beweis der Sätze 2.1 bis 2.13. Wegen $l_p^*\cong l_p^\dagger=l_q$ $\big(0<p<\infty;$ mit q wie in (2.1)$\big)$ und $l_\infty^\dagger\subset l_\infty^*$ folgt Satz 2.1 aus Satz 1.1, und da weiter $(e^{(k)})_k$ für $0<p<\infty$ eine Basis für l_p ist, folgen die Sätze 2.2 bzw. 2.3 aus Korollar 1.1 bzw. Satz 1.2.

Im folgenden sei stets $0<p<\infty$. Wegen $([\tilde{C}_\alpha]_\infty^p)^\dagger\subset([\tilde{C}_\alpha]_\infty^p)^*$ $\big($s. [1], (3.1)$\big)$ folgen die Sätze 2.4 und 2.5 aus Satz 1.1, und da $([\tilde{C}_\alpha]_0^p)^\dagger\cong([\tilde{C}_\alpha]_0^p)^*$ $\big($s. [1], Satz 3.1 (b)$\big)$ und $(e^{(k)})_k$ eine Basis für $[\tilde{C}_\alpha]_0^p$ ist $\big($s. [1], Satz 2.1 (b)$\big)$, folgen die Sätze 2.6 bzw. 2.7 aus Korollar 1.1 bzw. Satz 1.2.

Die Notwendigkeit der Bedingungen der Sätze 2.8 und 2.9 ist klar. Die Hinlänglichkeit der Bedingungen von Satz 2.9 ergibt sich wie folgt: Aus (ii) und (iii) folgt: zu jedem $k=1,2,\dots$ gibt es ein $a_k\in\mathbf{C}$ mit

$$\lim_{\mu\to\infty}\left\|\left((a_{nk}-a_k)^{(\mu)}(\beta;\,1)\right)_n\right\|_1=0, \tag{2.2}$$

und es gibt ein $a\in\mathbf{C}$ mit

$$\lim_{\mu\to\infty}\left\|\left((a-\sum_{k=1}^{\infty}a_{nk})^{(\mu)}(\beta;\,1)\right)_n\right\|_1=0. \tag{2.3}$$

Aus (i) und (ii) folgt weiter

$$(a_k)_k\in([\tilde{C}_\alpha]^p)^\dagger, \tag{2.4}$$

und es gibt ein $M\in\mathbf{R}$ mit

$$\sup_{\mu\in\mathbf{N}_0}\Big\{\max_{N_\mu\subset N^{(\mu)}}\Big\|\Big((1/A_{2^\mu-1}^\beta)\sum_{n\in N_\mu}A_{2^{\mu+1}-n}^{\beta-1}a_k\Big)_k\Big\|_{([\tilde{C}_\alpha]_\infty^p)^\dagger}\Big\}\leqq M. \tag{2.5}$$

Für alle $x\in[\tilde{C}_\alpha]^p$ und für alle $n\in\mathbf{N}$ schreiben wir

$$A_n(x)-l\Big(a-\sum_{k=1}^{\infty}a_k\Big)-\sum_{k=1}^{\infty}a_kx_k=\sum_{k=1}^{\infty}(a_{nk}-a_k)(x_k-l)+l\Big(\sum_{k=1}^{\infty}a_{nk}-a\Big), \tag{2.6}$$

wobei $l\in\mathbf{C}$ mit $\lim_{\nu\to\infty}\|(x-le)^{(\nu)}(\alpha;p)\|_p=0$, a_k aus (2.2) $(k=1,2,\dots)$ und a aus

(2.3). Wegen (i) existiert $A_n(x)$ für alle $n\in\mathbf{N}$, wegen (iii) existiert $\sum_{k=1}^{\infty} a_{nk}$ für alle $n\in\mathbf{N}$, und es ist $(\sum_{k=1}^{\infty} a_{nk}-a)_n\in[\tilde{C}_\beta]_0^1$, und wegen (2.4) existieren $\sum_{k=1}^{\infty} a_k x_k$ und wegen $le\in[\tilde{C}_\alpha]_0^p$ auch $\sum_{k=1}^{\infty} a_k$. Nach Satz 2.6 gilt wegen (2.2), (2.5) und (i) mit $\tilde{A}=(\tilde{a}_{nk})_{n,k}:=(a_{nk}-a_k)_{n,k}$:

$$\tilde{A}\in([\tilde{C}_\alpha]_0^p, [\tilde{C}_\beta]_0^1).$$

Da $x-le\in[\tilde{C}_\alpha]_0^p$ ist, existiert $\tilde{A}_n(x-le)=\sum_{k=1}^{\infty}(a_{nk}-a_k)(x_k-l)$ für alle $n\in\mathbf{N}$, und es gilt $(\tilde{A}_n(x))_n\in[\tilde{C}_\beta]_0^1$. Aus (2.6) folgt daher

$$\Big(A_n(x)-\big(l(a-\sum_{k=1}^{\infty} a_k)+\sum_{k=1}^{\infty} a_k x_k\big)\Big)_n\in[\tilde{C}_\beta]_0^1,$$

d. h. $(A_n(x))_n\in[\tilde{C}_\beta]^1$.

Die Hinlänglichkeit der Bedingungen von Satz 2.8 ergibt sich genau wie im Fall von Satz 2.9 mit $a_k=0$ $(k=1, 2, \ldots)$ und $a=0$.

Da $c_0^*\cong c_0^+=l_1$ und $(e^{(k)})_k$ eine Basis für c_0 ist, folgen die Sätze 2.10 bzw. 2.11 aus Korollar 1.1 bzw. Satz 1.2.

Die Notwendigkeit der Bedingungen der Sätze 2.12 und 2.13 ist klar. Die Hinlänglichkeit der Bedingungen der Sätze 2.13 und 2.12 ergibt sich ähnlich wie im Fall der Sätze 2.9 und 2.8.

Literatur

[1] E. Malkowsky, Toeplitz-Kriterien für Matrizenklassen stark limitierbarer Folgen, *Acta Sci. Math.*, **48** (1985), 297—313.

[2] I. J. Maddox, On Kuttner's Theorem, *J. London Math. Soc.*, 43 (1968), 285—290.

[3] I. J. Maddox, *Elements of Functional Analysis*, Cambridge University Press (Cambridge, 1970).

[4] A. Peyerimhoff, Über ein Lemma von Herrn H. C. Chow, *J. London Math. Soc.*, 32 (1957), 33—36.

[5] B. Kuttner, B. Thorpe, Matrix Transformations of Strongly Summable Series, *J. London Math. Soc.* (2), **11** (1975), 195—206.

MATHEMATISCHES INSTITUT
JUSTUS-LIEBIG-UNIVERSITÄT
IHERINGSTR. 6
6300 GIESSEN, BRD

Acta Sci. Math., 50 (1986), 133—142

Error bounds for certain classes of quintic splines

S. SALLAM

1. Introduction. Interpolation of functions by quintic splines has been extensively studied by many authors (see e.g. [5], [6] and [7]). Recently, SALLAM [4] has presented some new types of quintic splines where an analysis of their corresponding error bounds in L^2-norm was presented. Our object is to continue this study and obtain error bounds for these new types of interpolatory splines in L^∞-norm.

In Section 2, we developed some preliminary formulas of interpolatory quintic splines under different continuity requirements and different given data together with a lacunary interpolation by quintic splines of certain functions without using function values. Section 3 is devoted to studying various convergence results of the presented interpolatory splines, namely the error bounds in L^∞-norm, and it is shown that the order of convergence remains the same as if function values are considered.

2. Construction of some quintic splines. Let

$$\Delta:\ 0 = x_0 < x_1 < \ldots < x_{N+1} = 1$$

denote a partition of $I=[0, 1]$. Denote by $\mathcal{S}^{(l)}_{N,5}$, $l=2, 3$, the class of quintic splines $q_s(x)$ such that:

(i) $q_s(x) \in C^l(I)$;

(ii) $q_s(x)$ is a quintic in each $[x_i, x_{i+1}]$, $i=0(1)N$. We set $f^{(k)}(x_j)=f_l^k$ and $q_s^{(k)}(x_j)=q_j^k$ stand for $D^k f(x_j)$ and $D^k q_s(x_j)$, respectively. Further define $h_i= =x_{i+1}-x_i$ and $h=\max\limits_i h_i$.

We now discuss the possibility of interpolation of some given function by elements in $\mathcal{S}^{(l)}_{N,5}$. It is well known that the following procedures are well defined according to Theorems 1 and 2 (cf. [4]).

Received February 27, 1984.

Definition 1. Given numbers f_i, f_i^2, $i=0(1)N+1$ and f_i^1, $i=0, N+1$; there exists a unique quintic spline $\hat{q}_s(x)\in\mathscr{S}_{N,5}^{(3)}$ such that

$$\begin{aligned} \hat{q}_s(x_i) &= f_i, \\ \hat{q}_s^{(2)}(x_i) &= f_i^2, \quad i=0(1)N+1, \\ \hat{q}_s^{(1)}(x_i) &= f_i^1, \quad i=0, N+1, \end{aligned} \tag{2.1}$$

provided that $h_i > h_{i-1}$, for all i.

Definition 2. Given numbers f_i^1, f_i^2, $i=0(1)N+1$ and f_i, $i=0, N+1$, there exists a unique quintic spline $q_s^*(x)\in\mathscr{S}_{N,5}^{(3)}$ such that

$$\begin{aligned} \overset{*}{q}{}_s^{(1)}(x) &= f_i^1, \\ \overset{*}{q}{}_s^{(2)}(x) &= f_i^2, \quad i=0(1)N+1, \\ \overset{*}{q}_s(x_i) &= f_i, \quad i=0, N+1. \end{aligned} \tag{2.2}$$

We now turn to prove the following.

Theorem 1. *Given numbers f_i, f_i^3, $i=0(1)N+1$ and f_i^2, $i=0, N+1$, there exists a unique quintic spline $\bar{q}_s(x)\in\mathscr{S}_{N,5}^{(3)}$ such that*

$$\begin{aligned} \bar{q}_s(x_i) &= f_i, \\ \bar{q}_s^{(3)}(x_i) &= f_i^3, \quad i=0(1)N+1, \\ \bar{q}_s^{(2)}(x_i) &= f_i^2, \quad i=0, N+1. \end{aligned} \tag{2.3}$$

Theorem 2. *Given numbers f_i^2, f_i^3, $i=0(1)N+1$ and f_i, $i=0, N+1$, there exists a unique quintic spline $\tilde{q}_s(x)\in\mathscr{S}_{N,5}^{(3)}$ such that*

$$\begin{aligned} \tilde{q}_s^{(2)}(x_i) &= f_i^2, \\ \tilde{q}_s^{(3)}(x_i) &= f_i^3, \quad i=0(1)N+1, \\ \tilde{q}_s(x_i) &= f_i, \quad i=0, N+1. \end{aligned}$$

To prove the above theorems, we need the following well-known lemma (cf. [2]).

Lemma 1. *If $p(x)$ is a quintic on $[0, 1]$, then*

$$\begin{aligned} p(x) = p(0)B_0(1-x)+p(1)B_0(x)+p^{(2)}(0)B_1(1-x)+p^{(2)}(1)B_1(x)- \\ -p^{(3)}(0)B_2(1-x)+p^{(3)}(1)B_2(x) \end{aligned} \tag{2.5}$$

where

$$\begin{aligned} B_0(x) = x, \quad B_1(x) = (x^4-x^5)/10+(3/20)(x^4-x), \\ B_2(x) = (x^5-x^4)/20+(x-x^4)/30. \end{aligned} \tag{2.6}$$

Proof of Theorem 1. In order to establish Theorem 1, one can express $\bar{q}_s(x)$ in $[x_i, x_{i+1}]$ using (2.5) and $x=x_i+th_i$, $0\leqq t\leqq 1$, in the following form

$$\bar{q}_s(x)=f_iB_0(1-t)+f_{i+1}B_0(t)+h_i^2\bar{q}_i^2B_1(1-t)+h_i^2\bar{q}_{i+1}^2B_1(t)-\\ -h_i^3f_i^3B_2(1-t)+h_i^3f_{i+1}^3B_2(t). \tag{2.7}$$

Similarly, in $[x_{i-1}, x_i]$ $\bar{q}_s(x)$ may be represented as

$$\bar{q}_s(x)=f_{i-1}B_0(1-t)+f_iB_0(t)+h_{i-1}^2\bar{q}_{i-1}^2B_1(1-t)+h_{i-1}^2\bar{q}_i^2B_1(t)-\\ -h_{i-1}^3f_{i-1}^3B_2(1-t)+h_{i-1}^3f_i^3B_2(t). \tag{2.8}$$

Straightforward calculation shows that $\bar{q}_s(x)\in C^3(I)$ is equivalent to

(2.9)
$$(3/10)h_{i-1}\bar{q}_{i-1}^2+(7/10)(h_{i-1}+h_i)\bar{q}_i^2+(3/10)h_i\bar{q}_{i+1}^2=2h_{i-1}^{-1}f_{i-1}-2(h_{i-1}^{-1}+h_i^{-1})f_i+\\ +2h_i^{-1}f_{i+1}-(1/15)h_{i-1}^2f_{i-1}^3-(1/10)(h_i^2-h_{i-1}^2)f_i^3+(1/15)h_i^2f_{i+1}^3,\quad i=1(1)N.$$

In matrix form (2.9) can be written as $A\bar{\mathbf{q}}=\mathbf{b}$, where $A=(a_{ij})$ with

$$a_{ij}=\begin{cases}\nu_i=(3/7)h_i/(h_i+h_{i-1}), & i<j,\\ 1, & i=j,\\ \mu_i=(3/7)h_{i-1}/(h_i+h_{i-1}), & i>j.\end{cases} \tag{2.10}$$

Clearly A is a diagonally dominant matrix and hence the quintic spline is determined uniquely. In the case of uniform distribution, (2.9) becomes

$$(3/10)\bar{q}_{i-1}^2+(14/10)\bar{q}_i^2+(3/10)\bar{q}_{i+1}^2=2h^{-2}(f_{i-1}-2f_i+f_{i+1})-(h/15)(f_{i-1}^3-f_{i+1}^3).$$

Proof of Theorem 2. Similar to that of the previous theorem, the system (2.5) is now replaced by

(2.11)
$$-h_{i-1}^{-1}\tilde{q}_{i-1}+(h_{i-1}^{-1}+h_i^{-1})\tilde{q}_i-h_i^{-1}\tilde{q}_{i+1}=-(3/20)h_{i-1}f_{i-1}^2-(7/20)(h_{i-1}+h_i)f_i^2-\\ -(3/20)h_if_{i+1}^2-(1/30)h_{i-1}^2f_{i-1}^3+(1/20)(h_{i-1}^2-h_i^2)f_i^3+(1/30)h_i^2f_{i+1}^3,\quad i=1(1)N.$$

Again the matrix of coefficients is diagonally dominant. In the case of a uniform partition, (2.10) will take the form

$$\tilde{q}_{i-1}-2\tilde{q}_i+\tilde{q}_{i+1}=(3/20)h^2(f_{i-1}^2+(14/3)f_i^2+f_{i+1}^2)+(h^3/30)(f_{i-1}^3-f_{i+1}^3).$$

We now proceed to develop a new lacunary quintic spline, which interpolates the lacunary data (function values and third derivatives) midway between the knots and first derivatives at the knots. The problem is stated in the following theorem.

Theorem 3. *Given arbitrary numbers* f_i^1, $i=0(1)N+1$; $f^{(p)}(z_i)$, $i=0(1)N$, $p=0, 3$, $z_i=(1/2)(x_i+x_{i+1})$ *and* f_i, $i=0, N+1$, *then there exists a unique quintic* *spline* $Q(x) \in \mathscr{S}^{(2)}_{N,6}$, *such that*

$$(2.12)\qquad \begin{aligned} Q^{(1)}(x_i) &= f_i^1, & i &= 0(1)N+1,\\ Q^{(p)}(z_i) &= f^{(p)}(z_i), & i &= 0(1)N; \quad p = 0,3,\\ Q(x_i) &= f_i, & i &= 0, N+1, \end{aligned}$$

provided that $h_i > h_{i-1}$ *for all* i.

In order to establish Theorem 3, we need the following lemma (see [3]).

Lemma 2. *If* $p(x)$ *is quintic on* $[0, 1]$, *then*

$$(2.13)\qquad \begin{aligned} p(x) = {} & p(0)A_0(x)+p(1/2)A_1(x)+p(1)A_2(x)+p^{(1)}(0)A_3(x)+\\ & +p^{(1)}_{(1)}A_4(x)+p^{(3)}(1/2)A_5(x) \end{aligned}$$

where

$$\begin{aligned} & A_0(x) = 4x^5-18x^4+26x^3-13x^2+1, \quad A_1(x) = 16(x^4-2x^3+x^2),\\ & A_2(x) = A_0(1-x), \quad A_3(x) = 2x^5-7x^4+9x^3-5x^2+x,\\ & A_4(x) = -A_3(1-x), \quad A_5(x) = (1/6)(-2x^5+5x^4-4x^3+x^2). \end{aligned}$$

Proof of Theorem 3. In $[x_i, x_{i+1}]$ $Q(x)$ can be expressed as

(2.14)

$$Q(x) = Q_iA_0(t)+f(z_i)A_1(t)+Q_{i+1}A_2(t)+h_if_i^1A_3(t)+h_if_{i+1}^1A_4(t)+h_i^3f^{(3)}(z_i)A_5(t).$$

Similarly in $[x_{i-1}, x_i]$ $Q(x)$ will be

$$(2.15)\qquad \begin{aligned} Q(x) = {} & Q_{i-1}A_0(t)+f(z_{i-1})A_1(t)+Q_iA_2(t)+h_{i-1}f_{i-1}^1A_3(t)+\\ & +h_{i-1}f_i^1A_4(t)+h_{i-1}^3f^{(3)}(z_{i-1})A_5(t). \end{aligned}$$

In order that $Q(x) \in C^2(I)$, we obtain

$$(2.16)\qquad \begin{aligned} & 6h_{i-1}^{-2}Q_{i-1}+26(h_{i-1}^{-2}+h_i^{-2})Q_i-6h_i^{-2}Q_{i+1} = 32\big(h_{i-1}^{-2}f(z_{i-1})-\\ & -h_i^{-2}f(z_i)\big)+10(h_i^{-1}+h_{i-1}^{-1})f_i^1+(1/3)\big(h_if^{(3)}(z_i)-h_{i-1}f^{(3)}(z_{i-1})\big). \end{aligned}$$

In matrix form (2.16) can be written as $A\mathbf{Q}=\mathbf{b}$, where $A=(a_{ij})$ and

$$(2.17)\qquad a_{ij} = \begin{cases} \nu_i = -(3/13)h_{i-1}^2/(h_i^2-h_{i-1}^2), & i<j,\\ 1, & i=j,\\ \mu_i = (3/13)h_i^2/(h_i^2-h_{i-1}^2), & i>j. \end{cases}$$

It is clear that $\nu_i<0$ if $h_i>h_{i-1}$ for all i and hence A is nonsingular [1]. Thus $Q(x)$ is uniquely determined.

If the nodes are uniformly distributed on I, then (2.16) reduces to

$$Q_{i-1}-Q_{i+1} = (16/3)\big(f(z_{i-1})-f(z_i)\big)+(10/3)hf_i^1-(h^3/18)\big(f^{(3)}(z_{i-1})-f^{(3)}(z_i)\big).$$

The coefficient matrix is skew-symmetric and hence nonsingular if N is even.

3. Error bounds. The purpose of this section is to obtain error estimates for hese new types of interpolatory quintic and interpolation by lacunary quintic splines which introduced in [4] and in Section 2, in L^∞-norm. We follow the idea of Prasad and Varma [3].

Throughout $K_{r,l}$, $K'_{r,l}$, etc. denote generic constants independent of the functions considered and maximum mesh spacing h. However, these constants in general do depend in particular upon the order of various derivatives.

In the sequel we treat, for the sake of brevity, only estimates for error bounds of the quintic splines $\bar{q}_s(x)$ and $Q(x)$, respectively. Error bounds for other types can be handled in a similar manner. We begin with the following theorem.

Theorem 4. *Let Δ be an arbitrary partition of I. If $f\in C^l(I)$, $l=3(1)6$, then, for the unique quintic spline $\bar{q}_s(x)$ associated with f and satisfying* (2.3), *we have*

(3.1)
$$|\bar{q}_s^{(r)}(x)-f^{(r)}(x)| \leqq K_{r,l}h^{l-r}\|f^{(l)}\|_\infty+K'_{r,l}h^{l-r}\omega(f^{(l)}, h), \quad r = 0(1)3; \quad l = 3(1)5,$$

$$|\bar{q}_s^{(r)}(x)-f^{(r)}(x)| \leqq K_{r,6}h^{6-r}\|f^{(6)}\|_\infty, \quad r = 0(1)3; \quad l = 6, \tag{3.2}$$

where $\omega(f^{(l)}, h)$ denotes the modulus of continuity of $f^{(l)}$.

We shall prove Theorem 4 for $l=6$, the proof needs the following lemma.

Lemma 3. *Let $f\in C^6(I)$, then*

$$|\bar{q}_s^{(2)}(x_i)-f^{(2)}(x_i)| \leqq K_{2,6}h^4\|f^{(6)}\|_\infty. \tag{3.3}$$

Proof. The condition that $\bar{q}_s(x)\in C^3(I)$ is equivalent to the system (2.9). Using Taylor's formula, it can be shown that

(3.4)
$$\begin{aligned}(3/10)h_{i-1}f_{i-1}^2+(7/10)(h_{i-1}+h_i)f_i^2+(3/10)h_i f_{i+1}^2 &= 2h_{i-1}^{-1}f_{i-1}-2(h_{i-1}^{-1}+h_i^{-1})f_i+\\ &+2h_i^{-1}f_{i+1}-(1/15)h_{i-1}^2f_{i-1}^3-(1/10)(h_i^2-h_{i-1}^2)f_i^3+(1/15)h_i^2f_{i+1}-\\ &-(1/720)(h_i^5+h_{i-1}^5)f^{(6)}(\xi_i), \quad \xi_i\in(x_{i-1}, x_{i+1}), \quad i = 1(1)N.\end{aligned}$$

From (2.9) and (3.4), we deduce that

$$\begin{aligned}(3/10)h_{i-1}(\bar{q}_{i-1}^2-f_{i-1}^2)+(7/10)(h_i+h_{i-1})(\bar{q}_i^2-f_i^2)+(3/10)h_i(\bar{q}_{i+1}^2-f_{i+1}^2) &=\\ = (1/720)(h_{i-1}^5+h_i^5)f^{(6)}(\xi_i)&\end{aligned}$$

which can be written as $Ae=\mathbf{z}$ and the entries of A are given by (2.10), $e_i^2=\bar{q}_i^2-f_i^2$, $i=1(1)N$ and

$$z_i = (1/504)\big((h_{i-1}^5+h_i^5)/(h_{i-1}+h_i)\big)f^{(6)}(\xi_i).$$

Let $A=I+B$ where $\|B\|_\infty=3/7$, we have $\mathbf{e}^2=(I+B)^{-1}\mathbf{z}$. Hence $\|\mathbf{e}^2\|_\infty \leqq (7/4)\|\mathbf{z}\|_\infty$ i.e.,

$$\|\mathbf{e}^2\|_\infty \leqq K_{2,6} h^4 \|f^{(6)}\|_\infty,$$

where $K_{2,6}=1/288$. This completes the proof of Lemma 3.

It can be easily seen that the following identities are valid (cf. (2.6)):

$$\begin{aligned} &B_0(1-t)+B_0(t) = 1,\\ &B_0(t)+2B_1(1-t)+2B_1(t) = t^2,\\ &B_0(t)+6B_1(t)-6B_2(1-t)+6B_2(t) = t^3,\\ &B_0(t)+12B_1(t)+24B_2(t) = t^4,\\ &B_0(t)+20B_1(t)+60B_2(t) = t^5. \end{aligned} \tag{3.5}$$

Proof of Theorem 4. From (2.7), it follows that

$$h_i^3\bar{q}_s^{(3)}(x) = -h_i^2\bar{q}_i^2B_1'''(1-t)+h_i^2\bar{q}_{i+1}^2B_1'''(t)+h_i^3f_i^3B_2'''(1-t)+h_i^3f_{i+1}^3B_2'''(t), \tag{3.6}$$

where $B'''(z)=d^3B/dz^3$, which can be written as

$$h_i^3\bar{q}^{(3)}(x) = \lambda_i(t)+\mu_i(t) \tag{3.7}$$

where

$$\lambda_i(t) = -h_i^2(\bar{q}_i^2-f_i^2)B_1'''(1-t)+h_i^2(\bar{q}_{i+1}^2-f_{i+1}^2)B_1'''(t)$$

and

$$\mu_i(t) = -h_i^2f_i^2B_1'''(1-t)+h_i^2f_{i+1}^2B_1'''(t)+h_i^3f_i^2B_2'''(1-t)+h_i^3f_{i+1}^3B_2'''(t).$$

On using Lemma 3, we obtain

$$|\lambda_i(t)| \leqq K_{2,6}ah^4\|f^{(6)}\|_\infty, \tag{3.8}$$

where $a=\max\limits_{0\leqq z\leqq 1}|B_1'''(z)|$.

Using Taylor's expansion and the identities (3.5) together with

$$f^{(3)}(x) = \sum_{k=3}^{5}\left(f^{(k)}(x_i)/(k-3)!\right)(th_i)^{k-3}+\left(f^{(6)}(\eta_1)/5!\right)(th_i)^3, \tag{3.9}$$

it is not difficult to show that

$$\mu_i(t) = h_i^3f_i^3+h_i^4tf_i^4+(h_i^5/2)t^2f_i^5+(h_i^6/6!)[30f^{(6)}(\eta_2)B_1'''(t)+120f^{(6)}(\eta_3)B_2'''(t)]. \tag{3.10}$$

Thus, (3.9) and (3.10) yield

$$\mu_i(t) = h_i^3f^3(x)+(h_i^6/6!)\nu_i(t), \tag{3.11}$$

where

$$\nu_i(t) = 30f^{(6)}(\eta_2)B_1'''(t)+120f^{(6)}(\eta_3)B_2'''(t)-120t^3f^{(6)}(\eta_1). \tag{3.12}$$

Now combining (3.7) with (3.12), gives

$$h_i^3(\bar{q}_s^{(3)}(x)-f^{(3)}(x)) = \lambda_i(t)+(h_i^6/6!)\,v_i(t). \tag{3.13}$$

It is clear from (3.12) that

$$|v_i(t)| \leqq \{30|B_1'''(t)|+120|B_2'''(t)|+120|t^5|\} \max_{x\in I} |f^{(6)}(x)| \leqq C\|f^{(6)}\|_\infty. \tag{3.14}$$

Hence (3.13), (3.14) and (3.8) give

$$|\bar{q}_s^{(3)}(x)-f^{(3)}(x)| \leqq K_{3,6}h^3\|f^{(6)}\|_\infty.$$

This proves (3.2) for $r=3$. To prove (3.2) for $r=2$, observe that for $x_i \leqq x < x_{i+1}$ and using Lemma 3, it follows that

$$|\bar{q}_s^{(2)}(x)-f^{(2)}(x)| = \Big|\int_{x_i}^{x} (q_s^{(3)}(t)-f^{(3)}(t))\,dt+q_i^2-f_i^2\Big| \leqq K_{2,6}h^4\|f^{(6)}\|_\infty.$$

Since $\bar{q}_s(x)-f(x)$ vanishes at x_i and x_{i+1}, there exists a point ξ_i, $\xi_i\in(x_i, x_{i+1})$, such that $\bar{q}_s^{(1)}(\xi_i)=f^{(1)}(\xi_i)$. Hence

$$|\bar{q}_s^{(1)}(x)-f^{(1)}(x)| \leqq \int_{\xi_i}^{x} |\bar{q}_s^{(2)}(t)-f^{(2)}(t)|\,dt \leqq K_{1,6}h^5\|f^{(6)}\|_\infty.$$

This proves (3.2) for $r=1$. Formula (3.2), for $r=0$, follows immediately by using a similar argument. For $l=3(1)5$ the proof is analogous to that of Theorem 5.

We now turn to the derivation of error bounds for the interpolation error, $Q(x)-f(x)$, and its derivatives. We begin with the following main result.

Theorem 5. *Let $f\in C^l(I)$ and $h_i>\sqrt{8/5}\,h_{i-1}$ for all i or the partition be uniform. Then for the unique quintic spline $Q(x)$ associated with f and satisfying* (2.12), *we have*

$$|Q^{(r)}(x)-f^{(r)}(x)| \leqq K_{r,l}h^{l-r}\omega(f^{(l)}, h), \quad r=0(1)2; \quad l=3(1)5, \tag{3.15}$$

$$|Q^{(r)}(x)-f^{(r)}(x)| \leqq K_{r,6}h^{6-r}\|f^{(6)}\|_\infty+h^{6-r}\omega(f^{(6)}, h), \quad r=0(1)2; \quad l=6. \tag{3.16}$$

To prove Theorem 5, we need the following lemma

Lemma 4. *Let $f\in C^l(I)$ and $h_i>\sqrt{8/5}\,h_{i-1}$ for all i, then*

$$|Q(x_i)-f(x_i)| \leqq K_l(h_i^{l-1}h_{i-1}^{l-1}/(h_i^{l-2}-h_{i-1}^{l-2}))\,\omega(f^{(l)}, h), \quad l=3(1)5, \tag{3.17}$$

$$|Q(x_i)-f(x_i)| \leqq K_6h_i^2h_{i-1}^2(h_i^2+h_{i-1}^2)\|f^{(6)}\|_\infty, \quad l=6. \tag{3.18}$$

Proof. We will prove the lemma for $l=3$. The condition that $Q(x)\in C^2(I)$ is equivalent to the system (2.16). Using Taylor's expansion, it is easy to show that

$$\begin{aligned}(3.19)\qquad & 6h_{i-1}^{-2}(Q_{i-1}-f_{i-1})+26(h_{i-1}^{-2}-h_i^{-2})(Q_i-f_i)-6h_i^{-2}(Q_{i+1}-f_{i+1}) = \\ &= (2/3)h_i\big(f^{(3)}(\eta_{1,i})-f^{(3)}(\eta_{3,i})\big)+(2/3)h_{i-1}\big(f^{(3)}(\eta_{2,i})-f^{(3)}(\eta_{4,i})\big)+ \\ &+(h_i/3)\big(f^{(3)}(\eta_{1,i})-f^{(3)}(z_i)\big)+(h_{i-1}/3)\big(f^{(3)}(\eta_{2,i})-f^{(3)}(z_{i-1})\big),\end{aligned}$$

where $x_{i-1}<\eta_{2,i}<x_i$, $x_i<\eta_{1,i}<x_{i+1}$, $x_i<\eta_{3,i}<z_i$, $z_{i-1}<\eta_{4,i}<x_i$.

In matrix form (3.19) can be written as $A\mathbf{e}=\mathbf{z}$ with $e_i=Q(x_i)-f(x_i)$, $i=1(1)N$. Multiplying $A\mathbf{e}=\mathbf{z}$ by the diagonal matrix $D=(d_{ii})$, $d_{ii}=(1/26)h_i^2h_{i-1}^2/(h_i^2-h_{i-1}^2)$, the matrix DA will be $I+B$ with $\|B\|_\infty<1$ if $h_i>\sqrt{8/5}\,h_{i-1}$ for all i. Since

$$\|\mathbf{e}\| \leqq \|(I+B)^{-1}\|_\infty\|D\mathbf{z}\|_\infty \leqq \big(1/(1-\|B\|_\infty)\big)\|D\mathbf{z}\|_\infty.$$

It follows that

$$|Q(x_i)-f(x_i)| \leqq K_3\big(h_i^2h_{i-1}^2/(h_i-h_{i-1})\big)\omega(f^{(3)}, h)$$

with $K_3=5/13$. This proves (3.17) for $l=3$. The proof is similar in the other cases.

For equally spaced knots and N is even Lemma 4 will be modified as follows.

Lemma 5. *Let $f\in C^l(I)$, then*

$$(3.20)\qquad |Q(x_i)-f(x_i)| \leqq K_l h^l \omega(f^{(l)}, h), \quad l = 3(1)5.$$

Proof. We prove the lemma for $l=3$. The condition that $Q(x)\in C^2(I)$ and the partition be uniform is equivalent to (cf. (3.19))

$$\begin{aligned}&(Q_{i-1}-f_{i-1})-(Q_{i+1}-f_{i+1}) = (h^3/9)\big(f^{(3)}(\eta_{1,i})-f^{(3)}(\eta_{3,i})\big)+ \\ &+(h^3/9)\big(f^{(3)}(\eta_{2,i})-f^{(3)}(\eta_{4,i})\big)+(h^3/18)\big(f^{(3)}(\eta_{1,i})-f^{(3)}(z_i)\big)+ \\ &+(h^3/9)\big(f^{(3)}(\eta_{2,i})-f^{(3)}(z_{i-1})\big).\end{aligned}$$

Or, in the matrix form $M\mathbf{e}=\mathbf{z}$, where $M=(m_{ij})$ with

$$m_{ij} = \begin{cases} -1, & i<j \\ 0, & i=j \\ 1, & i>j. \end{cases}$$

Since $\|Mx\|_\infty\geqq 1$ for $\|x\|_\infty=1$, it follows that $\|M^{-1}\|_\infty\leqq 1$. Hence

$$\|\mathbf{e}\|_\infty \leqq \|M^{-1}\|_\infty\|\mathbf{z}\|_\infty \leqq K_3h^3\omega(f^{(3)}, h).$$

This completes the proof of Lemma 5.

It is well known that the following identities are valid (see [3]).

$$A_0(t)+A_1(t)+A_2(t) = 1,$$
$$A_1(t)+2A_2(t)+2A_3(t)+2A_4(t) = 2t,$$
$$(1/4)A_1(t)+A_2(t)+2A_4(t) = t^2,$$
$$(1/8)A_1(t)+A_2(t)+3A_4(t)+6A_5(t) = t^3,$$
$$(1/16)A_1(t)+A_2(t)+4A_4(t)+12A_5(t) = t^4.$$

Proof of Theorem 5. The proof will be carried out only for $l=3$. From (2.14), it follows that

$$h_i^2Q^{(2)}(x) = Q_iA_0''(t)+f(z_i)A_1''(t)+Q_{i+1}A_2''(t)+h_if_i^1A_3''(t)+ +h_if_{i+1}^1A_4''(t)+h_i^3f^{(3)}(z_i)A_5''(t) \tag{3.21}$$

which can be written as

$$h_i^2Q^{(2)}(x) = \lambda_i(t)+\mu_i(t), \tag{3.22}$$

where

$$\lambda_i(t) = \big(Q(x_i)-f(x_i)\big)A_0''(t)+\big(Q(x_{i+1})-f(x_{i+1})\big)A_2''(t),$$

and

$$\mu_i(t) = f(x_i)A_0''(t)+f(x_{i+1})A_2''(t)+f(z_i)A_1''(t)+h_if^{(1)}(x_i)A_3''(t)+ +h_if^{(1)}(x_{i+1})A_4''(t)+h_i^3f^{(3)}(z_i)A_5''(t).$$

Using Lemma 4, for $l=3$, we obtain

$$|\lambda_i(t)| \leqq K_3ah_i^2\left\{\frac{h_{i-1}^2}{h_i-h_{i-1}}+\frac{h_{i+1}^2}{h_{i+1}-h_i}\right\}\omega(f^{(3)},h) \tag{3.23}$$

where $a=\max\{l_0, l_2\}$, $l_r=\max\limits_{0\leqq t\leqq 1}|A_r''(t)|$.

Also it can be easily seen that

$$\mu_i(t) = h_i^2f^{(2)}(x)+(h_i^3/3!)\nu_i(t), \tag{3.24}$$

where

$$\nu_i(t) = (1/8)\big(f^{(3)}(\eta_{2,i})-f^{(3)}(z_i)\big)A_1''(t)+\big(f^{(3)}(\eta_{1,i})-f^{(3)}(\eta_{6,i})\big)A_2''(t)+ +3\big(f^{(3)}(\eta_{3,i})-f^{(3)}(\eta_{6,i})\big)A_4''(t)+6\big(f^{(3)}(z_i)-f^{(3)}(\eta_{6,i})\big)A_5''(t).$$

Consequently

$$|\nu_i(t)| \leqq a^*\omega(f^{(3)},h), \tag{3.25}$$

where $a^*=\max\{l_1, l_2, l_4, l_5\}$.

Combining (3.22) and (3.24), we have

$$h_i^2\big(Q^{(2)}(x)-f^{(2)}(x)\big) = \lambda_i(t)+(h_i^3/3!)\,\nu_i(t).$$

Thus

$$(3.26)\quad |Q^{(2)}(x)-f^{(2)}(x)| \leqq \left\{K_3 a \left(\frac{h_{i-1}^2}{h_i-h_{i-1}}+\frac{h_{i+1}^2}{h_{i+1}-h_i}\right)+\frac{a^*}{3!}h_i\right\}\omega(f^{(3)}, h) \leqq$$
$$\leqq K_{2,3}h\omega(f^{(3)}, h).$$

This proves (3.15) for $r=2$. The proof for $r=1$ and $r=0$ follows immediately using the fact that $Q^{(1)}(x)-f^{(1)}(x)$ vanishes at x_i and x_{i+1} and that $Q(x)-f(x)$ vanishes at $x=z_{i-1}$ and z_i.

Remark. It is worth noting that even in the absence of the function values at the mesh points, it is possible to construct quintic and lacunary quintic splines, which, for some cases, requires certain partition restrictions. Also the order of convergence achieved is almost the same when using function values as given data. We believe that this observed phenomenon may have some applications in physics and engineering problems as well.

References

[1] A. Bleyer and S. Sallam, Interpolation by cubic splines, *Periodica Poly.*, **22** (1978), 91—105.
[2] A. Meir and A. Sharma, Lacunary interpolation by splines, *SIAM J. Numer. Anal.*, **10** (1973), 423—442.
[3] J. Prasad and A. K. Varma, Lacunary interpolation by quintic splines, *SIAM J. Numer. Anal.*, **16** (1978), 1075—1079.
[4] S. Sallam, On interpolation by quintic splines, *Bull. Fac. Sci., Assiut University*, **11** (1982), 97—106.
[5] A. Sharma, Some poised and nonpoised problems of interpolation, *SIAM Rev.*, **14** (1972), 129—151.
[6] B. K. Swartz and R. Varga, A note on lacunary interpolation by splines, *SIAM J. Numer. Anal.*, **10** (1973), 443—447.
[7] B. K. Swartz and R. Varga, Error bounds for spline and *L*-spline interpolation, *J. Approx. Theory*, **6** (1972), 6—49.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF KUWAIT
P. O. BOX 5969
KUWAIT

Acta Sci. Math., 50 (1986), 143—157

Completely bounded maps and hypo-Dirichlet algebras

R. G. DOUGLAS and V. I. PAULSEN

1. An important result of VON NEUMANN [16] is the theorem that the closed unit disk $\mathbf{D}^-$ is a spectral set for every contraction T defined on a complex Hilbert space $\mathfrak{H}$. Sz.-NAGY showed in [25] that every contraction T possesses a unitary dilation U on some larger Hilbert space $\mathfrak{K}$, and thus obtained an alternate proof of this. The investigation of the relation of T and U forms the basis for [26]. In [24] Sz.-NAGY also proved that an invertible operator T satisfying $\sup\{\|T^k\|: k\in Z\}<\infty$ is similar to a unitary operator. His question of whether the assumption that $\sup\{\|T^k\|: k=1,2,\dots\}<\infty$, implies that T is similar to a contraction was shown to have a negative answer by FOGUEL [6]. HALMOS then reformulated the question [11] to ask if every polynomially bounded operator is similar to a contraction.

All of these results and questions can be reformulated to involve mappings from a function algebra A to the algebra $\mathfrak{L}(\mathfrak{H})$ of bounded operators on a Hilbert space $\mathfrak{H}$. Von Neumann's result concerns the extendibility of the mapping $p(z)\to p(T)$ to a contractive unital homomorphism from the disk algebra $A(\mathbf{D})$ to $\mathfrak{L}(\mathfrak{H})$. Sz.-Nagy's theorem shows that every such map dilates to a *-homomorphism from $C(\partial\mathbf{D})$ to $\mathfrak{L}(\mathfrak{K})$ for some Hilbert space $\mathfrak{K}$ containing $\mathfrak{H}$. Halmos' question asks whether every bounded unital homomorphism φ from $A(\mathbf{D})$ to $\mathfrak{L}(\mathfrak{H})$ is similar to a contractive unital homomorphism. ANDO [2] has shown that Sz.-Nagy's dilation theorem generalizes to the bidisk algebra $A(\mathbf{D}^2)$, and on the basis of PARROTT's example [17], one can show that this is false for $A(\mathbf{D}^n)$ for $n>2$ (cf. [26]). VAROPOULOS showed in [28] that the analogous result is also false for the ball algebra $A(\mathbf{B}^n)$ for $n>2$.

The above results form the core of dilation theory and the theory of spectral sets. Dilation theory is concerned with which linear maps from a function algebra to $\mathfrak{L}(\mathfrak{H})$ dilate to a representation of a self-adjoint algebra containing the function

Received September 14, 1983, and in revised form October 2, 1984.

The research of the authors was supported in part by grants from the National Science Foundation.

algebra, on some possibly larger Hilbert space. The theory of spectral sets is concerned with determining when a particular set Ω is spectral for an operator T and if it is, deciding whether or not T possesses a normal dilation whose spectrum is contained in $\partial\Omega$, i.e., a $\partial\Omega$-normal dilation. That is, deciding whether or not the induced contractive homomorphism, $r(z)\to r(T)$, on the uniform closure $R(\Omega)$ of the rational functions with poles off Ω, dilates to $C(\partial\Omega)$.

Foiaş in [7] and then independently Berger [4] and Lebow [13] (cf. Sarason [21]) showed that if $R(\Omega)$ is a Dirichlet algebra on $\partial\Omega$, then all unital contractive homomorphisms on $R(\Omega)$ dilate to $C(\partial\Omega)$.

In this paper we investigate problems for mappings from function algebras defined for finitely connected domains in $\mathbf{C}$. Such domains yield algebras $R(\Omega)$ which are hypo-Dirichlet. We show, in Section 2, that operators having such domains as spectral sets are similar to operators which have a normal dilation. We prove, more generally, that all unital contractive homomorphisms of hypo-Dirichlet algebras are similar to homomorphisms that dilate.

While these results still leave open the question of whether or not operators having such domains as spectral sets have normal dilations, other results suggest that perhaps this is the wrong question. The search for criteria which insure that a set is spectral has been in some ways unsuccessful. Shields [33] remarks that for an invertible operator T, the annulus with radii determined by the norms of T and T^{-1} need not be spectral for T. In [14], Misra shows that this occurs even when T is a 2 by 2 matrix. However, in Section 3, we show that such an operator is always similar to an operator for which the original annulus is spectral, and which moreover possesses a normal dilation. Thus, criteria for a set to be a spectral set, up to similarity, can be more readily obtainable. An analogous phenomenon occurs for all multiply connected domains—adding a similarity allows one to consider simpler domains.

In [3], Arveson reformulated dilation theory and generalized some of the previous results. If φ is a linear mapping from a function algebra A to $\mathfrak{L}(\mathfrak{H})$, then there is a natural linear mapping φ_n from $M_n(A)$ to $M_n(\mathfrak{L}(\mathfrak{H}))\cong\mathfrak{L}(\mathfrak{H}\otimes\mathbf{C}^n)$, where $M_n(A)$ denotes the algebra of $n\times n$ matrices with entries from A, defined by applying φ entrywise. If φ is bounded, then so is φ_n and $\|\varphi\|\leqq\|\varphi_2\|\leqq\|\varphi_3\|\leqq\dots$, but it is not necessarily true that $\|\varphi_n\|=\|\varphi\|$ or even that the sequence of norms is uniformly bounded. When it is, the map is said to be *completely bounded* and we set $\|\varphi\|_{cb}=\sup\{\|\varphi_n\|: n=1, 2, \dots\}$. The map is said to be *n-contractive* if $\|\varphi_n\|\leqq 1$ and *completely contractive* if it is n-contractive for all n.

Generalizing earlier work of Stinespring [23], Arveson showed in [3] that a unital contractive map is dilatable if and only if it is completely contractive. Moreover, he showed that all unital contractive maps on Dirichlet algebras are completely contractive, thus generalizing the results of Sz.-Nagy and Foiaş—Berger—

Lebow. Ando's result shows that all unital contractive homomorphisms on the bidisk algebra are completely contractive, but the question for unital contractive linear maps is open.

In Section 4, we show that Arveson's result on Dirichlet algebras does not extend to hypo-Dirichlet algebras. We give an example for the annulus algebra of a unital contraction which is not even a 2-contraction, although we leave unanswered the question for homomorphisms. (See Note added in proof.)

A necessary condition that a homomorphism be similar to a dilatable homomorphism is that it be completely bounded. Extending the work of ARVESON [3], and HADWIN [10], the second author has shown that this is also sufficient [18] (see also, [9], [31], [32]). In fact, if φ is a unital completely bounded homomorphism then there is always a similarity S, with $\|S\|\cdot\|S^{-1}\|=\|\varphi\|_{cb}$ such that $S\varphi S^{-1}$ is completely contractive [19]. Thus, in particular an operator T is similar to a contraction if and only if the mapping $p(z)\to p(T)$ extends to be completely bounded on $A(\mathbf{D})$. The problem of whether T being polynomially bounded implies that the resulting map is completely bounded is still unsolved.

2. A uniformly closed unital subalgebra A of $C(X)$ is said to be hypo-Dirichlet of codimension n, if the closure $\mathfrak{S}$ of $A+\bar{A}$ is of finite codimension n in $C(X)$. A Dirichlet algebra is obviously hypo-Dirichlet as is the algebra $R(\Omega)$ generated by the rational function on a finitely connected domain in $\mathbf{C}$. We consider in this section the problem of dilating unital contractive mappings from a hypo-Dirichlet algebra A to $\mathfrak{L}(\mathfrak{H})$. In Section 4 we present an example showing that Arveson's result doesn't extend even to the annulus algebra although our example is not a homomorphism. We show instead that for $\varphi\colon A\to\mathfrak{L}(\mathfrak{H})$ a unital contraction, one has $\|\varphi\|_{cb}\leqq 2n+1$, where n is the codimension of $\mathfrak{S}$ in $C(X)$. In particular, this implies that if $\varphi\colon A\to\mathfrak{L}(\mathfrak{H})$ is a unital contractive homomorphism, then there exists an invertible operator S such that $\|S^{-1}\|\,\|S\|\leqq 2n+1$ and the mapping $\varphi_1\colon A\to\mathfrak{L}(\mathfrak{H})$ defined by

$$\varphi_1(f)=S^{-1}\varphi(f)S$$

is dilatable to $C(X)$.

We begin with the following lemma about bases in finite dimensional Banach spaces.

Lemma 2.1. *Let A be a unital C^*-algebra and let $1\in\mathfrak{S}\subseteqq A$ be a self-adjoint subspace of codimension n. Then for every $\varepsilon>0$ there exists a positive map $\varrho\colon A\to\mathfrak{S}$ and a positive linear functional s on $\mathfrak{S}$ such that $\|\varrho\|\leqq n+1+\varepsilon$, $\|s\|\leqq n+\varepsilon$, and*

$$\varrho(f)=f+s(f)\cdot 1 \quad \text{for } f \text{ in } \mathfrak{S}.$$

Proof. Let $\pi\colon A\to A/\mathfrak{S}$ be the quotient map. It is not difficult to show that there exist self-adjoint linear functionals $l_1', l_2', \ldots, l_n'$ on $A/\mathfrak{S}$ and self-adjoint elements $h'_1, h'_2, \ldots, h'_n$ in $A/\mathfrak{S}$ which form a basis, such that $\|l'_i\|=\|h'_i\|=1$ and $l'_i(h'_j)=\delta_{ij}$, the Kronecker delta. Let $h_1, h_2, \ldots, h_n$ be self-adjoint elements in A such that $\pi(h_i)=h'_i$ and $\|h_i\|\leqq 1+\varepsilon/n$. Also, let $l_i=l'_i\circ\pi$, so that $\mathfrak{S}=\{f\in A\colon l_i(f)=0, i=1, 2, \ldots, n\}$ with $\|l_i\|=1$ and $l_i(h_j)=\delta_{ij}$.

By [20, 1.14.3] we can write $l_i=p_i-q_i$, where $\|p_i\|\leqq 1$ $\|q_i\|\leqq 1$ and $\|p_i+q_i\|\leqq 1$ with p_i and q_i positive linear functionals in A. For every g in A

$$f = g - \sum_{i=1}^{n} l_i(g)h_i,$$

is in $\mathfrak{S}$. We define a positive map $\varrho\colon A\to\mathfrak{S}$ by

$$\varrho(g) = g + \sum_{i=1}^{n} q_i(g)(\|h_i\|+h_i) + \sum_{i=1}^{n} p_i(g)(\|h_i\|-h_i).$$

Since each of the three expressions defines a positive map, we need only check that the range of ϱ is contained in $\mathfrak{S}$. An easy calculation shows for $g=f+\sum_{i=1}^{n} l_i(g)h_i$ that $\varrho(g)=f+\sum_{i=1}^{n}(p_i(g)+q_i(g))\|h_i\|$. If we set $s(g)=\sum_{i=1}^{n}(p_i(g)+q_i(g))\|h_i\|$, then

$$\|s\| \leqq \sum_{i=1}^{n} \|p_i+q_i\|\,\|h_i\| \leqq n+\varepsilon$$

and $\varrho(f)=f+s(f)\cdot 1$ for f in $\mathfrak{S}$. Finally, since ϱ is positive we have $\|\varrho\|=\|\varrho(1)\|=\|1+s(1)\|\leqq n+1+\varepsilon$.

Theorem 2.2. *If $A\subseteqq C(X)$ is a hypo-Dirichlet algebra of codimension n, and $\varphi\colon A\to\mathfrak{L}(\mathfrak{H})$ is a unital contraction, then $\|\varphi\|_{cb}\leqq 2n+1$.*

Proof. Fix $\varepsilon>0$, let $\mathfrak{S}$ be the closure of $A+\bar{A}$ in $C(X)$ and let $\varrho\colon C(X)\to\mathfrak{S}$ and s be as in the previous Lemma. If we extend φ to $\tilde{\varphi}\colon\mathfrak{S}\to\mathfrak{L}(\mathfrak{H})$ by $\tilde{\varphi}(f+\bar{g})=\varphi(f)+\varphi(g)^*$, then $\tilde{\varphi}$ will be positive by [3, pp. 152—153]. Thus $\tilde{\varphi}\circ\varrho\colon C(X)\to\mathfrak{L}(\mathfrak{H})$ is positive and hence completely positive [23]. Finally for f in A, we have

$$\varphi(f) = \tilde{\varphi}\circ\varrho(f) - s(f)\cdot 1_{\mathfrak{H}}$$

so that,

$$\|\varphi\|_{cb} \leqq \|\tilde{\varphi}\circ\varrho\|_{cb} + \|s\|_{cb} = \|\tilde{\varphi}\circ\varrho(1)\| + \|s(1)\| \leqq 2n+1+2\varepsilon,$$

since $\tilde{\varphi}\circ\varrho$ and s are completely positive (see [3, Proposition 1.2.10]).

Corollary 2.3. *If $A\subseteqq C(X)$ is a hypo-Dirichlet algebra of codimension n, and $\varphi\colon A\to\mathfrak{L}(\mathfrak{H})$ is a unital contractive homomorphism, then φ is similar to a homomorphism that dilates to $C(X)$. Furthermore, the similarity S may be chosen such that $\|S\|\cdot\|S^{-1}\|\leqq 2n+1$.*

Corollary 2.4 *Let Ω be a spectral set for T in $\mathfrak{L}(\mathfrak{H})$ with $R(\Omega)$ a hypo-Dirichlet subalgebra of $C(\partial\Omega)$ of codimension n. Then there exists an invertible operator S on $\mathfrak{H}$ with $\|S\|\cdot\|S^{-1}\|\leqq 2n+1$ such that $S^{-1}TS$ has a $\partial\Omega$-normal dilation.*

We are unable to determine whether these results are the best one can do even in the case of the annulus.

3. As we mentioned in the first section the prototypical example of a hypo-Dirichlet algebra is $R(\Omega)$ for Ω the closure of a finitely connected domain in $\mathbf{C}$. As might be expected a more natural proof of Corollary 2.4 is possible in this case which has other consequences that we explore. We make no attempt at working with the most general domains possible.

A compact subset Ω of $\mathbf{C}$ is said to be a K-spectral set for T in $\mathfrak{L}(\mathfrak{H})$ if $\sigma(T)\subseteq\Omega$ and $\|f(T)\|\leqq K\|f\|$ for f in $R(\Omega)$, where $\|f\|$ denotes the supremum norm on $\partial\Omega$. We call Ω a *complete K-spectral set* for T if, in addition,

$$\|(f_{i,j}(T))\| \leqq K\|(f_{i,j})\|$$

for all matrices $(f_{i,j})$ in $M_n(R(\Omega))$, where $\|(f_{i,j})\|$ denotes the supremum of the matrix norm on $\partial\Omega$ and $(f_{i,j}(T))$ denotes the operator in $M_n(\mathfrak{L}(\mathfrak{H}))\cong\mathfrak{L}(\mathfrak{H}\otimes\mathbf{C}^n)$. For $K=1$ we obtain the usual notion of *spectral* or *complete spectral set.*

We will need also to consider certain unbounded subsets of $\mathbf{C}$. If Ω is a closed unbounded subset of $\mathbf{C}$, then we let $R(\Omega)$ denote the function algebra obtained from the uniform closure of the rational functions on Ω with poles off Ω and which vanish at infinity. We shall call such an Ω a *(complete) K-spectral set* for T provided that the above inequalities hold for all f in $R(\Omega)$ ($(f_{i,j})$ in $M_n(R(\Omega))$).

We shall call a compact subset Ω of $\mathbf{C}$ a *Dirichlet algebra domain* if $\{\mathrm{Re}\,(f)\colon f\in R(\Omega)\}$ is uniformly dense in the algebra of real-valued continuous functions on $\partial\Omega$. More generally, we shall call an arbitrary closed subset Ω a *Dirichlet algebra domain* if there is a λ in $C\setminus\Omega$ such that the compact set

$$\Omega_\lambda^{-1} = \{(z-\lambda)^{-1}\colon z\in\Omega\}\cup\{0\}$$

is a Dirichlet algebra domain.

We record the following facts about these "unbounded" Dirichlet domains for further use.

Proposition 3.1. *Let Ω be a closed subset of* $\mathbf{C}$ *containing a neighborhood of infinity and let λ be in* $\mathbf{C}\setminus\Omega$.

(i) *If Ω is a (complete) K-spectral set for T, then Ω_λ^{-1} is a (complete)* $(2K+1)$-*spectral set for* $(T-\lambda)^{-1}$.

(ii) *If Ω_λ^{-1} is a (complete) K-spectral set for* $(T-\lambda)^{-1}$, *then Ω is a (complete) K-spectral set for T.*

(iii) *The set* Ω *is a spectral set for* T *if and only if* Ω_λ^{-1} *is a spectral set for* $(T-\lambda)^{-1}$.

(iv) *If* Ω *is a Dirichlet algebra domain, then* Ω *is a spectral set for* T *if and only if* Ω *is a complete spectral set for* T.

Proof. If Ω is a K-spectral set for T, then for every f in $R(\Omega_\lambda^{-1})$, with $f(0)=0$, $g(z)=f((z-\lambda)^{-1})$ is in $R(\Omega)$ and hence,

$$\|f((T-\lambda)^{-1})\| = \|g(T)\| \leqq K\|g\| = K\|f\|.$$

Now, if f is an arbitrary element of $R(\Omega_\lambda^{-1})$, then,

$$\|f((T-\lambda)^{-1})\| \leqq \|f((T-\lambda)^{-1})-f(0)I\|+\|f(0)I\| \leqq (2K+1)\|f\|.$$

The proofs of the "complete" case and of (ii) are similar. If Ω is a spectral set for T, and f is in $R(\Omega_\lambda^{-1})$, with $\|f\|\leqq 1$, let $x=f(0)$ and let $\varphi_x(z)=(z-x)/(1-\bar{x}z)$ be the conformal mapping of the disc into itself that carries x to 0. We have that $\|\varphi_x \circ f\|<1$, and since $\varphi_x\circ f(0)=0$, by the calculations above,

$$\|\varphi_x \circ f((T-\lambda)^{-1})\| < 1.$$

But now by von Neumann's inequality,

$$\|f((T-\lambda)^{-1})\| = \|\varphi_{-x} \circ \varphi_x \circ f((T-\lambda)^{-1})\| < 1.$$

This conformal mapping technique was introduced by WILLIAMS [30]. (iv) follows easily from the case of bounded domain using (ii).

We shall call a compact subset Ω of $\mathbf{C}$ *decomposable* if $\Omega=\bigcap \Omega_j$ (possibly infinite), Ω_j is a Dirichlet algebra domain and there is a constant K such that for any n and f in $M_n(R(\Omega))$, we can write $f=\sum f_j$ (norm convergent) with f_j in $M_n(R(\Omega_j))$ and

$$\sum \|f_j\| \leqq K\|f\|.$$

We shall let K_Ω denote the minimum of such constants K. We show the usefulness of this notion after establishing that nice finitely connected subsets of $\mathbf{C}$ are decomposable.

Proposition 3.2. *If* Ω *is a compact subset of* $\mathbf{C}$ *such that* $\mathbf{C}\setminus\Omega$ *consists of finitely many open connected subsets whose boundaries are disjoint and each is a rectifiable simple closed curve, then* Ω *is decomposable.*

Proof. Let $\mathbf{C}\setminus\Omega=\Delta_0\cup\Delta_1\cup\ldots\cup\Delta_n$ with Δ_0 the connected component of infinity and let $\Gamma_i=\partial\Delta_i$ be given a parametrization such that $\Gamma=\bigcup\Gamma_i$ has winding number 1 for every point in the interior of Ω and 0 for every point in $\mathbf{C}\setminus\Omega$. Let $\Omega_i=\mathbf{C}\setminus\Delta_i$ so that Ω_i is a Dirichlet algebra domain and Ω_0 is bounded.

For λ in the interior of Ω and f in $M_k(R(\Omega))$, let

$$f_j(\lambda) = (1/2\pi i) \int_{\Gamma_j} (f(z)/(z-\lambda))\, dz$$

so that $f(\lambda)=\sum_{i=0}^{n} f_i(\lambda)$ for λ in Ω. Note that f_j clearly extends to a continuous function on Γ_i for $i\neq j$ and hence to a continuous function in Γ_j by letting $f_j(\lambda)= =f(\lambda)-\sum_{i\neq j} f_i(\lambda)$. Moreover, f_0 is analytic on the interior of Ω_0, continuous on Ω_0 and hence in $M_k(R(\Omega_0))$ [8, Chapter VIII, 8.4]. Similarly, for $i\neq 0$, f_i is in $M_k(R(\Omega_i))$.

Now for fixed λ in Γ_i, $i\neq j$,

$$\|f_j(\lambda)\| \leqq d(\Gamma_i, \Gamma_j)^{-1} \cdot |\Gamma_j| \cdot \|f\|$$

where $|\Gamma_j|$ is the length of Γ_j and $d(\Gamma_i, \Gamma_j)$ denotes the usual distance between Γ_i and Γ_j. Finally, for λ in Γ_j, we have

$$\|f_j(\lambda)\| \leqq \|f\| + \sum_{i\neq j} \|f_i(\lambda)\| \leqq \Big(1+\sum_{i\neq j} d(\Gamma_i, \Gamma_j)^{-1} \cdot |\Gamma_i|\Big)\|f\|$$

and thus $\sum \|f_i\| \leqq K\|f\|$ which completes the proof.

In [14] an example is given of an operator T for which $\{z\colon |z|\leqq R_1\}$ and $\{z\colon |z|\geqq R_0\}$ are both spectral sets for T, while the annulus $\{z\colon R_0\leqq |z|\leqq R_1\}$ is not a spectral set for T. The use of decomposability shows that no such example can exist if we consider K-spectral sets instead.

Theorem 3.3. *Let Ω be decomposable with $\Omega=\bigcap_{j\geqq 1} \Omega_j$. If each Ω_j is a (complete) K_j-spectral set for T and $\sup_j K_j<\infty$, then there exists a K such that Ω is a (complete) K-spectral set for T.*

Moreover, if each Ω_j is a spectral set for T, then Ω is a complete K_Ω-spectral set for T.

Proof. Let f be in $M_k(R(\Omega))$ and write $f=\sum_j f_j$ with f_j in $M_k(R(\Omega_j))$. If we set $M=\sup_j K_j$, then

$$\|f(T)\| \leqq \sum_j \|f_j(T)\| \leqq MK_\Omega \|f\|$$

from which the first result follows.

If each Ω_j is spectral for T, then since each Ω_j is a Dirichlet algebra domain, it is completely spectral by Proposition 3.1. Hence for f in $M_k(R(\Omega))$ we have

$$\|f(T)\| \leqq \sum_j \|f_j(T)\| \leqq K_\Omega \|f\|.$$

Corollary 3.4. *If $\Omega=\bigcap\Omega_j$ is decomposable and each Ω_j is spectral for T, then there exists an invertible operator S with $\|S^{-1}\|\cdot\|S\|\leqq K_\Omega$ such that $S^{-1}TS$ has a $\partial\Omega$-normal dilation.*

Proof. Apply the Theorem together with [18].

We note that if Ω is decomposable and spectral for T then the hypotheses of Corollary 3.4 are certainly met. Thus we obtain that for Ω a "nice" n-holed domain whose holes are separated, then up to similarity every spectral operator is dilatable. The hypothesis that the holes of Ω are separated is necessary only for the proof we gave and many domains not satisfying this are decomposable. For example, the domain $\Omega=\Omega_0\cap\Omega_1$, where $\Omega_0=\{z\colon |z|\leqq 2\}\cap\{z\colon |z+1|\geqq 1\}$ and $\Omega_1=\{z\colon |z-3/4|\geqq 1/4\}\cap\{z\colon |z-5/4|\geqq 1/4\}$ is decomposable with Dirichlet algebra domains Ω_0 and Ω_1. By combining these techniques with the Mlak decomposition theorem [15], one can extend considerably the class of sets for which spectral implies similar to a dilatable operator. However, it is not clear at this time whether or not these techniques yield any sets Ω for which $R(\Omega)$ is not hypo-Dirichlet. It would also be interesting to have more particular information on the decomposability constant K_Ω versus the value $2n+1$ in the hypo-Dirichlet case.

We conclude this section with a different application of the notion of decomposability.

Corollary 3.5. *If T is a bounded invertible operator on $\mathfrak{H}$ for which there exist invertible operators S_1 and S_2 satisfying*

$$\|S_1^{-1}TS_1\|\leqq\alpha \quad \text{and} \quad \|S_2^{-1}T^{-1}S_2\|\leqq\beta,$$

then there exists an invertible operator S such that both

$$\|S^{-1}TS\|\leqq\alpha \quad \text{and} \quad \|S^{-1}T^{-1}S\|\leqq\beta.$$

Proof. If $\alpha\beta\leqq 1$, then $\sigma(T)\subseteq\{z\colon |z|\leqq\alpha\}$ and $\sigma(T^{-1})\subseteq\{z\colon |z|\leqq\beta\}$ which implies that

$$\sigma(T)\subseteq\{z\colon \beta^{-1}\leqq|z|\leqq\alpha\} \quad \text{and hence} \quad \alpha\beta=1.$$

Setting $R=\alpha^{-1}T$ we have that $\|R^n\|\leqq\|S_1^{-1}\|\,\|S_1\|$ and $\|R^{-n}\|\leqq\|S_2^{-1}\|\,\|S_2\|$ for $n>0$. Therefore, by the result of Sz.-Nagy [24] there exists an invertible operator S such that $S^{-1}RS$ is unitary which implies that $\|S^{-1}TS\|\leqq\alpha$ and $\|S^{-1}T^{-1}S\|\leqq\beta$.

If $\alpha\beta>1$, then $\Omega_0=\{z\colon |z|\leqq\alpha\}$ is a spectral set for $S_1^{-1}TS_1$, and hence a complete spectral set since Ω_0 is a Dirichlet algebra domain. This implies for f in $M_k(R(\Omega_0))$ that

$$\|f(T)\|=\|(S_1\otimes 1_k)f(S_1^{-1}TS_1)(S_1\otimes 1_k)^{-1}\|\leqq\|S_1^{-1}\|\,\|S_1\|\,\|f\|$$

and so Ω_0 is a complete $(\|S_1^{-1}\|\,\|S_1\|)$-spectral set for T. Similarly, $\Omega_1=\{z\colon |z|\geqq\beta^{-1}\}$ is a complete $(\|S_2^{-1}\|\,\|S_2\|)$-spectral set for T. Therefore, by Theorem 3.3, $\Omega=\Omega_0\cap\Omega_1$, is a complete K-spectral set for T for some K. Again applying [18], we have that there exists an invertible operator S with $\|S^{-1}\|\,\|S\|\leqq K$ and such that Ω is a complete spectral set for $S^{-1}TS$. In particular, this implies that $\|S^{-1}TS\|\leqq\alpha$ and $\|S^{-1}T^{-1}S\|\leqq\beta$.

One difficulty mitigating the usefulness of results on dilating completely bounded maps is the lack of uniqueness. However, even in the class of operators for which the annulus is a complete spectral set and hence dilatable, there is no uniqueness to the normal dilation as was observed in [1]. In the next section we pursue a more restrictive notion of dilation on the annulus for which uniqueness persists.

4. In this section we show that even for the annulus algebra $R(A)$, the analogue of Arveson's result for Dirichlet algebras is false. Namely, we show that there exists a unital contractive linear map $\varphi\colon R(A)\to\mathfrak{L}(\mathfrak{H})$ which is not completely contractive. We remind the reader that in this case the closure of the real parts of the rational functions on A with poles off A is a subspace of codimension one in the algebra of continuous real-valued functions on ∂A.

We begin by tying together some results on completely positive maps. The equivalence of statements (1) to (5) is certainly known [3], while (6) is based on an idea of Choi [5]. Recall that a linear map φ defined from a self-adjoint subspace $\mathfrak{S}$ of a C^*-algebra A to $\mathfrak{L}(\mathfrak{H})$ is said to be *completely positive* if $\varphi_n\colon M_n(\mathfrak{S})\to\mathfrak{L}(\mathfrak{H}\otimes\mathbf{C}_n)$ is positive for $n=1, 2, 3, \ldots$.

Proposition 4.1. *Let A be a unital C^*-algebra, $1\in\mathfrak{R}\subset A$ a subalgebra and let $\mathfrak{S}=\mathfrak{R}+\mathfrak{R}^*$ (or its closure). Then the following are equivalent:*

(1) *Every unital contraction $\varphi\colon\mathfrak{R}\to\mathfrak{L}(\mathfrak{H})$ is a complete contraction,*

(2) *Every unital, positive $\varphi\colon\mathfrak{S}\to\mathfrak{L}(\mathfrak{H})$ is completely positive,*

(3) *Every positive $\varphi\colon\mathfrak{S}\to\mathfrak{L}(\mathfrak{H})$ is completely positive,*

(4) *For all n, every unital positive $\varphi\colon\mathfrak{S}\to M_n$ is completely positive,*

(5) *For all n, every positive $\varphi\colon\mathfrak{S}\to M_n$ is completely positive, and*

(6) *For all n, the convex hull $h(\mathfrak{S}^+\otimes M_n^+)$ of $\{f\otimes p\colon f\in\mathfrak{S}^+, p\in M_n^+\}$ is norm dense in $(\mathfrak{S}\otimes M_n)^+$.*

Proof. The fact that (2) implies (1) follows from [3, Proposition 1.2.8 and Theorem 1.2.9]. The equivalence of (2) and (4) follows from restricting $\mathfrak{H}$ and φ to finite dimensional subspaces. Similarly (3) and (5) are equivalent. The equivalence of (4) and (5) follows by considering $\varphi(1)=R$, restricting to the subspace where it is invertible and replacing φ by $R^{-1/2}\varphi(\cdot)R^{-1/2}$. Since $\mathfrak{R}$ is a subalgebra, by von Neumann's inequality, unital positive maps on $\mathfrak{S}$ are contractive on $\mathfrak{R}$ and so (1) implies (2).

We now establish the equivalence of (5) and (6). To do this we need to recall that there is a one to one correspondence between linear maps from $\mathfrak{S}$ into M_n

and linear functionals on $\mathfrak{S}\otimes M_n$, such that under this correspondence a linear map is completely positive if and only if the associated linear functional is positive [12] (see also [22]). For $\varphi\colon \mathfrak{S}\to M_n$ a linear map, the corresponding linear functional s_φ is given by $s_\varphi((f_{i,j}))=\sum \varphi(f_{i,j})_{i,j}$, where the outer subscript indicates taking the (i,j)-th entry of the matrix.

Now if φ is positive, f is in $\mathfrak{S}^+$, p is in M_n^+; and we set $Q=\varphi(f)$, then

$$s_\varphi(f\otimes p) = \sum p_{i,j}q_{i,j}$$

which is easily recognizable as the sum of the entries of the Schur product $(p_{i,j}q_{i,j})$ of p and Q. Since the Schur product of positive matrices is positive and since the sum of the entries of a positive matrix is positive, we have that $s_\varphi(f\otimes p)\geqq 0$. Hence, if φ is positive, then s_φ is positive on every thing in the convex hull $h(\mathfrak{S}^+\otimes M_n^+)$ of $\{f\otimes p\colon f\in\mathfrak{S}^+,\ p\in M_n^+\}$. Thus, when $h(\mathfrak{S}^+\otimes M_n^+)$ is dense in $(\mathfrak{S}\otimes M_n)^+$, s_φ will be positive and consequently, φ will be completely positive.

Conversely, if $h(\mathfrak{S}^+\otimes M_n^+)$ is not dense, then choose a linear functional s which is positive on $h(\mathfrak{S}^+\otimes M_n^+)$ but negative somewhere on $(\mathfrak{S}\otimes M_n)^+$. If $\varphi_s\colon \mathfrak{S}\to M_n$ is the associated linear map from $\mathfrak{S}$ to M_n, then φ_s is not completely positive. However, if f is in $\mathfrak{S}^+$ and $x=(x_1, x_2, \dots, x_n)$ is a vector in $\mathbf{C}^n$, then

$$\langle \varphi_s(f)x, x\rangle = \sum \varphi_s(f)_{i,j}\cdot x_j\bar{x}_i = s(f\otimes p) \geqq 0,$$

where $p=(\bar{x}_i x_j)$ is a matrix in M_n^+. Thus φ is positive but not completely positive.

One advantage of this result is that it allows one to replace problems involving the selection of measures, by problems involving the approximation of matrix-valued functions. As an example of this we present a new proof of the well-known result that every positive map on $C(X)$ is completely positive [23].

Corollary 4.2. *Every positive map $\varphi\colon C(X)\to\mathfrak{L}(\mathfrak{H})$ is completely positive.*

Proof. It is enough by the Proposition to show that p in $(C(X)\otimes M_n)^+$ can be approximated by a sum $\sum f_i p_i$, when f_i is in $C(X)^+$ and p_i is in M_n^+. To this end fix $\varepsilon>0$ and choose a finite open covering $\{U_i\}$ of X such that for x, y in U_j we have $\|p(x)-p(y)\|<\varepsilon$. Let $f_1, \dots, f_n$ be a positive partition of unity of X subordinate to the U_j's. Fix points x_j in U_j, set $p_j=p(x_j)$ and note that

$$\|p-\sum f_i p_j\| < \varepsilon.$$

We remark that if one were interested in the question of when a unital k-contractive (k-positive) map is completely contractive (positive), there is an appropriate analogue of Proposition 4.1. One just replaces contractive by k-contractive and positive by k-positive in (1) to (4) and $h(\mathfrak{S}^+\otimes M_n^+)$ by a bigger set (a set which contains, in particular, all elements of the form, $f\oplus 0$ with f in $(\mathfrak{S}\otimes M_k^+)$). The standard results, such as, the fact that every k-positive map of $C(X)\otimes M_k$ is completely positive can then be proved along the same lines as that for Corollary 4.2. Further analogous results for C^*-algebras having only finite dimensional irreducible representations of bounded dimension can also be obtained.

We turn our attention now to the annulus $A=\{z\colon R_0\leqq|z|\leqq R_1\}$, where $0<R_0<R_1$ and let $\mathfrak{S}$ denote the closure of $R(A)+\overline{R(A)}$ in $C(\partial A)$. It was proved in [29] that $\mathfrak{S}$ is a subspace of $C(\partial A)$ of codimension one and, in fact, that $C(\partial A)$ is the span of $\mathfrak{S}$ and $\log|z|$. From these facts, it is clear that $C(\partial A)$ is also the span of $\mathfrak{S}$ and $h(z)$, where we set

$$h(z)=\begin{cases}1, & |z|=R_1,\\ 0, & |z|=R_0.\end{cases}$$

We also need another characterization of $\mathfrak{S}$. For this we define positive linear functions s_j, $j=0, 1$, on $C(\partial A)$ by

$$s_j(f)=(1/2\pi)\int_0^{2\pi} f(R_j\cdot e^{i\theta})\,d\theta,\quad j=0,1.$$

Note that if $f(z)=\sum_{j=-N}^{N} a_j z^j$ is a Laurent polynomial, then $s_0(f)=s_1(f)=a_0$. Thus $s_0(f)=s_1(f)$ for all f in $\mathfrak{S}$. Since $\mathfrak{S}$ has codimension one in $C(\partial A)$, it follows that

$$\mathfrak{S}=\{f\in C(\partial A)\colon s_0(f)=s_1(f)\}.$$

Note also that $s_j(h)=j$, $j=0, 1$. We fix these notations for the remainder of this section.

We begin with the negative result.

Theorem 4.3. *There exists a unital contractive map* $\varphi\colon R(A)\to M_2$ *which is not completely contractive.*

Proof. By virtue of Proposition 4.1, it is sufficient to construct an F in $(\mathfrak{S}\otimes M_2)^+$ which is not approximable by elements of $h(\mathfrak{S}^+\otimes M_2^+)$.

We define $F=(f_{i,j})$ as follows:

$$f_{12}(R_1e^{i\theta})=f_{21}(R_1e^{i\theta})=0,$$

$$f_{12}(R_0e^{i\theta})=f_{21}(R_0e^{i\theta})=\begin{cases}2\theta/\pi, & 0\leqq\theta\leqq\pi/2,\\ 2(\pi-\theta)/\pi, & \pi/2\leqq\theta\leqq 3\pi/2,\\ 2(\theta-2\pi)/\pi, & 3\pi/2\leqq\theta\leqq 2\pi,\end{cases}$$

$$f_{11}(R_0e^{i\theta})=f_{22}(R_0e^{i\theta})=1,$$

$$f_{11}(R_1e^{i\theta})=f_{22}(R_1e^{-i\theta})=\begin{cases}8\theta/\pi, & 0\leqq\theta\leqq\pi/2,\\ 8(\pi-\theta)/\pi, & \pi/2\leqq\theta\leqq\pi,\\ 0, & \pi\leqq\theta\leqq 2\pi.\end{cases}$$

Note that $s_0(f_{i,j})=s_1(f_{i,j})$ and that F is positive at each point so that F is in $(\mathfrak{S}\otimes M_2)^+$.

Fix $\varepsilon>0$ and suppose there exist g_l in $\mathfrak{S}^+$ and P_l in M_2^+, $l=1, 2, \dots, n$ such that $\|F-\sum_{l=1}^{n} g_l P_l\|<\varepsilon$. Let

$$a_{l,j} = (1/\pi)\int_0^{\pi} g_l(R_j e^{i\theta})\,d\theta, \quad j=0,1; \quad l=1,2,\dots,n,$$

and let

$$b_{l,j} = (1/\pi)\int_{\pi}^{2\pi} g_l(R_j e^{i\theta})\,d\theta, \quad j=0,1; \quad l=1,2,\dots,n.$$

Note that $a_{l,0}+b_{l,0}=2s_0(g_l)=2s_1(g_l)=a_{l,1}+b_{l,1}$.

Also let

$$A_j=(1/\pi)\int_0^{\pi} F(R_j e^{i\theta})\,d\theta, \quad B_j=(1/\pi)\int_{\pi}^{2\pi} F(R_j e^{i\theta})\,d\theta, \quad j=0,1,$$

so that

$$A_1=\begin{pmatrix}2 & 0\\ 0 & 0\end{pmatrix}, \quad A_0=\begin{pmatrix}1 & 1/2\\ 1/2 & 1\end{pmatrix}, \quad B_1=\begin{pmatrix}0 & 0\\ 0 & 2\end{pmatrix} \quad \text{and} \quad B_0=\begin{pmatrix}1 & -1/2\\ -1/2 & 1\end{pmatrix}.$$

Note that

$$\Big\|A_j-\sum_l a_{l,j}P_l\Big\|<\varepsilon \quad \text{and} \quad \Big\|B_j-\sum_l b_{l,j}P_l\Big\|<\varepsilon, \quad j=0,1,$$

since all are integrated over the same intervals.

Let $\sum_p$ denote summation over those indices for which $\operatorname{Re}((P_l)_{12})\geqq 0$. We then have that

$$(\textstyle\sum_p a_{l,1}P_l)_{11} \leqq (\sum a_{l,1}P_l)_{11} < (A_1)_{11}+\varepsilon = 2+\varepsilon$$

and

$$(\textstyle\sum_p a_{l,1}P_l)_{22} \leqq (\sum a_{l,1}P_l)_{22} < (A_1)_{22}+\varepsilon = \varepsilon.$$

Since $\det(\sum_p a_{l,1}P_l)\geqq 0$, we have that

$$\operatorname{Re}(\textstyle\sum_p(a_{l,1}P_l)_{12}) \leqq |(\sum_p a_{l,1}P_l)_{12}| < \sqrt{(2+\varepsilon)\varepsilon}.$$

Similarly, we find that

$$\operatorname{Re}(\textstyle\sum_p(b_{l,1}P_l)_{12}) \leqq |(\sum_p b_{l,1}P_l)_{12}| \leqq \sqrt{(2+\varepsilon)\varepsilon}.$$

However, since $0\leqq a_{l,0}\leqq a_{l,1}+b_{l,1}$, we see that

$$\operatorname{Re}((\textstyle\sum a_{l,0}P_l)_{12}) \leqq \operatorname{Re}((\sum_p a_{l,0}P_l)_{12}) \leqq \operatorname{Re}((\sum_p (a_{l,1}+b_{l,1})P_l)_{12}) < 2\sqrt{(2+\varepsilon)\varepsilon}.$$

Now, it is clear that for ε sufficiently small, this contradicts the fact that

$$\varepsilon > \Big\|A_0-\sum a_{l,0}P_l\Big\| \geqq |(A_0-\textstyle\sum a_{l,0}P_l)_{12}| = |1/2-(\sum a_{l,0}P_l)_{12}|.$$

This contradiction shows that $h(\mathfrak{S}^+\otimes M_2^+)$ is not dense in $(\mathfrak{S}\otimes M_2)^+$ which completes the proof of the Theorem.

Before continuing let us make a remark. If φ is a unital contractive linear map from $\mathfrak{S}$ to M_2 which is not completely contractive it is clear by Proposition 4.1 that it is not two-contractive so that $\|\varphi_2\|>1$. It may be true for the annulus algebra $R(A)$ that a unital two contractive linear φ must be completely contractive. Although the above techniques can be used in principle to resolve this problem, we have been unable to do it.

We now turn our attention to a completely positive map from $C(\partial A)$ to $\mathfrak{S}$ related to that given in Lemma 2.1 which in the case of the annulus we can describe explicitly.

Theorem 4.4. *The map* $\varrho\colon C(\partial A)\to\mathfrak{S}$ *defined by*

$$\varrho(g) = [g+s_0(g)h+s_1(g)(1-h)]/2$$

is a unital, completely positive map with range $\mathfrak{S}$.

Proof. The proof is straightforward.

The explicit nature of ϱ allows us to construct dilations, something like the ϱ-dilations of [23]. Note in particular, that $\varrho(z^n)=z^n/2$ for $n\neq 0$.

Theorem 4.5. *Let* $A=\{z\colon R_0\leqq|z|\leqq R_1\}$ *be a spectral set for* T *in* $\mathfrak{L}(\mathfrak{H})$. *Then there is a Hilbert space* $\mathfrak{K}\supseteqq\mathfrak{H}$ *and a normal operator* N *on* $\mathfrak{K}$ *with* $\sigma(N)\subseteqq\partial A$ *such that*

$$T^n = 2P_{\mathfrak{H}}N^n|\mathfrak{H},\quad n\neq 0,\quad \text{and}\quad (R_0^2+R_1^2)/2 = P_{\mathfrak{H}}N^*N|\mathfrak{H}.$$

Moreover, if $\mathfrak{K}$ *is the smallest reducing subspace for* N *containing* $\mathfrak{H}$, *then* N *is unique up to unitary equivalence.*

Proof. Let ψ be the unital contractive homomorphism from $R(A)$ to $\mathfrak{L}(\mathfrak{H})$ defined by $\psi(f)=f(T)$ and $\tilde{\psi}$ the positive extension of ψ from $\mathfrak{S}$ to $\mathfrak{L}(\mathfrak{H})$. If we let $\varphi=\tilde{\psi}\circ\varrho$, then φ is a unital positive map from $C(\partial A)$ to $\mathfrak{L}(\mathfrak{H})$ and hence is a unital completely positive map. Applying Stinespring's Theorem to φ we obtain a Hilbert space $\mathfrak{K}\supseteqq\mathfrak{H}$ and a *-homomorphism $\pi\colon C(\partial A)\to\mathfrak{L}(\mathfrak{H})$ such that $\varphi(\cdot)=P_{\mathfrak{H}}\pi(\cdot)|\mathfrak{H}$. Let $N=\pi(z)$ so that N is normal, $\sigma(N)\subseteqq\partial A$, and

$$2P_{\mathfrak{H}}N^n|\mathfrak{H} = 2\tilde{\psi}\circ\varrho(z^n) = \psi(z^n) = T^n.$$

Also since $|z|^2=(R_1^2-R_0^2)h(z)+R_0^2$ we have that

$$P_{\mathfrak{H}}N^*N|\mathfrak{H} = \tilde{\psi}\circ\varrho\big((R_1^2-R_0^2)h(z)+R_0^2\big) = (R_0^2+R_1^2)/2.$$

Finally, the uniqueness statement comes from the uniqueness of a minimal Stinespring representation.

Note added in proof. Recently Jim Agler in "Rational Dilation on an Annulus" (Ann. Math., 121 (1985), 537—564) has proven that every operator for which the annulus A is a spectral set, possesses a ∂A-normal dilation.

References

[1] M. B. ABRAHAMSE and R. G. DOUGLAS, Operators on multiply-connected domains, *Proc. Royal Irish Acad*, 74A (1974), 135—141.

[2] T. ANDO, On a pair of commutative contractions, *Acta Sci. Math.*, 24 (1963), 88—90.

[3] W. B. ARVESON, Subalgebras of C^*-algebras, *Acta Math.*, 123 (1969), 141—224.

[4] C. A. BERGER, *Normal dilations*, Doctoral dissertation, Cornell University, 1963.

[5] M.-D. CHOI, Positive linear maps, in: *Operator algebras and applications*, Proc. Symp. Pure Math., Vol. 38, Part II, Amer. Math. Soc. (Providence, R. I., 1982); pp. 583—590.

[6] S. R. FOGUEL, A counterexample to a problem of Sz.-Nagy, *Proc. Amer. Math. Soc.*, 15 (1964), 788—790.

[7] C. FOIAŞ, Certaines applications des ensembles spectraux. I. Mesure harmonique-spectrale, *Stud. Cerc. Mat.*, 10 (1959), 365—401.

[8] T. W. GAMELIN, *Uniform Algebras*, Prentice Hall (Englewood Cliffs, N. J., 1969).

[9] U. HAAGERUP, Solution of the similarity problem for cyclic representations of C^*-algebras, *Ann. Math.*, 118 (1983), 215—240.

[10] D. HADWIN, Dilations and Hahn decompositions for linear maps, *Canad. J. Math.*, 33 (1981), 826—839.

[11] P. R. HALMOS, Ten problems in Hilbert space, *Bull. Amer. Math. Soc.*, 70 (1970), 887—933.

[12] C. LANCE, On nuclear C^*-algebras, *J. Functional Analysis*, 12 (1973), 157—176.

[13] A. LEBOW, On von Neumann's theory of spectral sets, *J. Math. Anal. Appl.* 7 (1963), 64—90.

[14] G. MISRA, Curvature inequalities and extremal properties of bundle shifts, *J. Operator Theory*, 11 (1984), 305—318.

[15] W. MLAK, Decompositions and extensions of operator valued representations of function algebras, *Acta Sci. Math.*, 30 (1969), 181—193.

[16] J. VON NEUMANN, Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes, *Math. Nachr.*, 4 (1931), 258—281.

[17]. S. PARROTT, Unitary dilations for commuting contractions, *Pacific J. Math.*, 34 (1973), 481—490.

[18] V. I. PAULSEN, Every completely polynomially bounded operator is similar to a contraction, *J. Functional Analysis*, 55 (1984), 1—17.

[19] V. I. PAULSEN, Completely bounded homomorphisms of operator algebras, *Proc. Amer. Math. Soc.*, 92 (1984), 225—228.

[20] S. SAKAI, *C^*-algebras and W^*-algebras*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 60, Springer-Verlag (Berlin, 1971).

[21] D. E. SARASON, On spectral sets having connected complement, *Acta Sci. Math.*, 26 (1965), 289—299.

[22] R. SMITH and J. WARD, Matrix ranges for Hilbert space operators, *Amer. J. Math.*, 102 (1980), 621—631.

[23] W. F. STINESPRING, Positive functions on C^*-algebras *Proc. Amer. Math. Soc.* 6 (1955), 211—216.

[24] B. Sz.-Nagy, On uniformly bounded linear transformations in Hilbert space, *Acta Sci. Math.*, **11** (1947), 152—157.

[25] B. Sz.-Nagy, Sur les contractions de l'espace de Hilbert, *Acta Sci. Math.*, **15** (1953), 87—92.

[26] B. Sz.-Nagy and C. Foiaş, *Harmonic Analysis of Operators on Hilbert Space*, American Elsevier (New York, 1970).

[27] M. Takesaki, *Theory of Operator Algebras. I*, Springer-Verlag (New York, 1979).

[28] N. Th. Varopoulos, On an inequality of von Neumann and application of the metric theory of tensor products to operators theory, *J. Functional Analysis*, **16** (1974), 83—100.

[29] J. L. Walsh, The approximation of harmonic functions by harmonic polynomials and by harmonic rational functions, *Bull. Amer. Math. Soc.*, **35** (1929), 499—544.

[30] J. P. Williams, Minimal spectral sets of compact operators, *Acta Sci. Math.*, **28** (1967), 93—106.

[31] G. Wittstock, Ein operatorwertiger Hahn—Banach Satz, *J. Functional Analysis*, **40** (1981), 127—150.

[32] G. Wittstock, *Extensions of completely bounded C*-module homomorphisms*, Monographs and Studies in Math., Vol. 18, Pitman (Boston, 1984); pp. 238—250.

[33] A. Shields, Weighted shift operators and analytic function theory, in: *Topics in operator theory*, Mathematical Surveys, Vol. 13, Amer. Math. Soc. (Providence, R. I., 1974); pp. 49—128.

(R.G.D.)
DEPARTMENT OF MATHEMATICS
STATE UNIVERSITY OF NEW YORK AT STONY BROOK
STONY BROOK, NEW YORK 11 794, U.S.A.

(V.I.P.)
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF HOUSTON
HOUSTON, TEXAS 77 004, U. S. A.

Acta Sci. Math., 50 (1986), 159—167

A generalization of a theorem of Dieudonné for k-triangular set functions

ENDRE PAP

1. Introduction

Although non-additive set functions occur frequently in mathematics (semi-variations of measures with values in abstract spaces, outer measures, capacities, etc.), just recently are they studied in detail. In recent years several authors considered non-additive set functions.

As it is well-known, the Nikodym boundedness theorem for measures in general fails for algebras of sets (see Example 5., DIESTEL, UHL [2], p. 18). But there are uniform boundedness theorems in which the initial boundedness conditions are on some subfamilies of a given σ-algebra; those subfamilies may not be σ-algebras. A famous theorem of DIEUDONNÉ [3] states that for compact metric spaces the pointwise boundedness of a family of Borel regular measures on open sets implies its uniform boundedness on all Borel sets. We shall generalize Dieudonné's theorem on a wider class of set functions. The class of finitely additive regular Borel set functions gives nothing new, because each finitely additive regular Borel set functions (also in the case of vector measures) is necessarily countably additive (KUPKA [7]).

We take in this paper a wider class of real valued set functions, the so-called k-triangular set functions ([5], [6]). We prove a generalized Dieudonné type theorem for this class of set functions. Using some modifications we obtain also a generalization of Dieudonné's theorem for semigroup valued set functions.

Received February 17, 1984.

2. k-triangular set functions

Let T be a locally compact space and $\mathcal{S}$ a class of subsets of T such that $\emptyset \in \mathcal{S}$. First some definitions.

Definition 1 (Dinculeanu [4], p. 303). A set function $\mu: \mathcal{S} \to R$ is said to be *regular* if for every $A \in \mathcal{S}$ and every $\varepsilon > 0$ there exist a compact set $K \subset A$ and an open set $G \supset A$ such that for every set $A' \in \mathcal{S}$, $K \subset A' \subset G$, we have

$$|\mu(A) - \mu(A')| < \varepsilon.$$

Definition 2. A set function $\mu: \mathcal{S} \to R_+$ is said to be *k-triangular* with $k \in (0, +\infty)$ if for every $A, B \in \mathcal{S}$, such that $A \cap B = \emptyset$ and $A \cup B \in \mathcal{S}$, we have

$$\mu(A) - k\mu(B) \leqq \mu(A \cup B) \leqq \mu(A) + k\mu(B)$$

and $\mu(\emptyset) = 0$.

The following theorem is important for further characterization of set functions which are both regular and triangular.

Theorem 1. *Let $\mathcal{S}$ be a ring of subsets of T. If a set function $\mu: \mathcal{S} \to R$ is regular and superadditive, i.e.*

$$\mu(A \cup B) \geqq \mu(A) + \mu(B) \quad \textit{for every} \quad A, B \in \mathcal{S}, \quad A \cap B = \emptyset,$$

then it satisfies the following condition

(R) *For every $A \in \mathcal{S}$ and every number $\varepsilon > 0$ there exist a compact set $K \subset A$ and an open set $G \supset A$ such that for every set $B \in \mathcal{S}$ with $B \subset G \setminus K$ we have* $|\mu(B)| < \varepsilon$.

Proof. It is enough to adapt the proof of Proposition 1 on the page 304 in [4].

Corollary 1. *If a set function $\mu: \mathcal{S} \to R$ ($\mathcal{S}$ is a ring), $\mu(\emptyset) = 0$, has regular variation, where the variation $|\mu|$ is defined in the usual way, i.e.*

$$|\mu|(E) := \sup_{\pi} \sum_{A \in \pi} |\mu(A)| \quad (E \in \mathcal{S})$$

and the supremum is taken over all partitions π of E into a finite number of pairwise disjoint members of $\mathcal{S}$, then μ satisfies condition (R).

Proof. Since $|\mu|$ is superadditive ([4], p. 34), we can apply Theorem 1 for $|\mu|$. Then the inequality $\mu \leqq |\mu|$ implies our statement.

Definition 3. A set function $\mu: \mathcal{S} \to R$ is said to be *exhaustive* whenever given a sequence (E_n) of pairwise disjoint members of $\mathcal{S}$, $\lim_{n \to \infty} \mu(E_n) = 0$.

It is obvious that a k-triangular set function μ with regular variation is itself regular.

3. Uniform boundedness theorem

We take from now on for the class $\mathscr{S}$ the collection $\mathscr{B}$ of all Borel sets of a Hausdorff locally compact topological space T. Now we formulate the main theorem.

Theorem 2. *Let $\mathscr{M}$ be a family of k-triangular set functions defined on $\mathscr{B}$ with regular variations. If the set* $\{\mu(O);\, \mu\in\mathscr{M}\}$ *is bounded for every open set O, then*

$$\{\mu(B);\ \mu\in\mathscr{M},\ B\in\mathscr{B}\}$$

is a bounded set.

Remark 1. We shall assume in the following proofs that T is a compact Hausdorff space. Namely, we can replace T with an Alexandrov one point ω compactification $T\cup\{\omega\}$, taking $\mu(\omega)=0$ $(\mu\in\mathscr{M})$.

We easily obtain the following

Corollary 2. *Let $\mathscr{M}$ be a family of regular scalar measures defined on $\mathscr{B}$. If the set* $\{|\mu(O)|;\, \mu\in\mathscr{M}\}$ *is bounded for every open set O, then*

$$\{|\mu(B)|;\ \mu\in\mathscr{M},\ B\in\mathscr{B}\}$$

is a bounded set.

Proof. Let $v(B):=|\mu(B)|$ $(B\in\mathscr{B},\ \mu\in\mathscr{M})$. It is obvious that the family $\mathscr{F}$ of all such set functions v satisfies the conditions of Theorem 2 (by Proposition 24 from [4], p. 319, $|v|=|\mu|$ is also regular). So we apply Theorem 2.

In the proof of Theorem 2 we need two lemmas.

Lemma 1. *Let μ be a k-triangular set function defined on $\mathscr{B}$ with regular variation. Then μ is k-σ-subadditive on each sequence of disjoint open sets (O_n), i.e.*

$$\mu\Big(\bigcup_{j=1}^{\infty} O_j\Big) \leqq k \sum_{j=1}^{\infty} \mu(O_j).$$

Proof of Lemma 1. First we shall prove that μ is order continuous on open sets, i.e. for each sequence (U_n) of open sets such that $U_j\supset U_{j+1}$ $(j\in N)$ and $\bigcap_{j=1}^{\infty} U_j=\emptyset$, we have

$$\lim_{j\to\infty} \mu(U_j) = 0.$$

For each $\varepsilon>0$ there exists a sequence of compact sets (K_n) such that $K_j\subset U_j$ and

$$|\mu|(U_j\setminus K_j) < \begin{cases} \varepsilon/2^j & \text{for } 0<k\leqq 1 \\ \varepsilon/2^j k & \text{for } k>1 \end{cases} \quad (j\in N). \tag{1}$$

Then there exists $n_0\in N$ such that $\bigcap_{j=1}^{n} K_j=\emptyset$ for all $n\geqq n_0$. Let $n\geqq n_0$. Then we have

$$\mu(U_n) = \mu(U_n \setminus \bigcap_{j=1}^{n} K_j) = \mu(\bigcup_{j=1}^{n} (U_n \setminus K_j)) \leqq |\mu|(\bigcup_{j=1}^{n} (U_n \setminus K_j)).$$

Hence, since $|\mu|$ is k-subadditive (i.e. $|\mu|(A \cup B) \leqq |\mu|(A) + k|\mu|(B)$ for every pair A, B of not necessarily disjoint sets from $\mathscr{B}$, see [4], pp. 35—36 and p. 16) and non-decreasing, we obtain by (1) for $k \geqq 1$ (for $0 < k < 1$ we take $k = 1$)

$$\mu(U_n) \leqq k \sum_{j=1}^{n} |\mu|(U_j \setminus K_j) < \varepsilon$$

for all $n \geqq n_0$. Now, let (O_n) be a sequence of disjoint open sets. Then we have

$$\mu(\bigcup_{j=1}^{\infty} O_j) \leqq k \sum_{j=1}^{n} \mu(O_j) + \mu(\bigcup_{j=n+1}^{\infty} O_j).$$

Taking $n \to \infty$ we obtain

$$\mu(\bigcup_{j=1}^{\infty} O_j) \leqq k \sum_{j=1}^{\infty} \mu(O_j).$$

The following lemma is given by C. SWARTZ [12] as an extract from the elementary proof of the Antosik—Miskusiński diagonal theorem [1].

Lemma 2. *Let X be a Banach space. If $x_{ij} \in X$ $(i, j \in N)$ such that*

$$\lim_{j \to \infty} x_{ij} = 0 \quad (i \in N), \quad \lim_{i \to \infty} x_{ij} = 0 \quad (j \in N)$$

and $\|x_{ii}\| \geqq \delta > 0$ $(i \in N)$, then there exist a sequence (i_n) of natural numbers and a sequence (ε_n) of positive real numbers such that

$$\left\|\sum_{k=1}^{n-1} x_{i_n i_k}\right\| = (1/2 - \varepsilon_n)\|x_{i_n i_n}\|, \quad \|x_{i_n i_{n+q}}\| < 2^{-q}\varepsilon_n \|x_{i_n i_n}\|$$

(in [12] δ is instead of $\|x_{i_n i_n}\|$).

Proof of Theorem 2. Firstly, let us suppose that $k = 1$. It suffices to prove that every point in T belongs to an open set O on which holds

$$\sup\{\mu(A)\colon A \subset O \quad (A \in \mathscr{B}),\ \mu \in \mathscr{M}\} < \infty. \tag{2}$$

Suppose that this is not true. Then there exists a point $x \in T$ such that (2) does not hold for every open set O such that $x \in O$. We shall prove that there exists a sequence of pairwise disjoint open sets (E_n) and a sequence (μ_n) from $\mathscr{M}$ such that $\mu_i(E_i) > i$ $(i \in N)$. For any open set O such that $x \in O$ there exists a Borel set $B \subset O$ and $\mu_1 \in \mathscr{M}$ such that

$$\mu_1(B) > 4 + 2 \sup_{\mu \in \mathscr{M}} \mu(\{x\}). \tag{3}$$

It is easy to prove that the preceding supremum is finite. Since μ_1 has regular variation, by Corollary 1 there exists a compact set $K\subset B$ and an open set $O'\subset O$, $B\subset O'$ such that $\mu_1(B')<1$ for each $B'\subset O'\setminus K$. We have by the subadditivity of μ_1

$$\mu_1(K)+\mu_1(B\setminus K) \geqq \mu_1(B).$$

Using the preceding inequality, the inequality $\mu_1(B\setminus K)<1$ and (3), we obtain

$$\mu_1(K) > 3+2 \sup_{\mu\in\mathcal{M}} \mu(\{x\}).$$

Let $K_1=K\cup\{x\}$. Then the last inequality implies (directly for $x\in K$) by the triangularity of μ_1 (for $x\notin K$) that

$$\mu(K_1) > 3+ \sup_{\mu\in\mathcal{M}} \mu(\{x\}).$$

By the regularity of μ_1 there exists an open set U such that $O\supset U\supset K_1$ and $\mu_1(B'')<1$ for every $B''\subset U\setminus K_1$. The preceding inequality together with the inequality

$$\mu_1(U) \geqq \mu_1(K_1)-\mu_1(U\setminus K_1)$$

implies

$$\mu_1(U) > 2+ \sup_{\mu\in\mathcal{M}} \mu(\{x\}). \tag{4}$$

Again by the regularity of μ_1 there exists an open set W such that $\{x\}\subset W\subset U$ and

$$\mu_1(B''') < 1 \tag{5}$$

for every $B'''\subset W\setminus\{x\}$.

Let H be an open set such that $x\in H\subset \bar{H}\subset W$ (where $\bar{H}$ is the closure of the set H). Then we have

$$\mu_1(\bar{H}) \leqq \sup_{A\subset \bar{H}\setminus\{x\}} \mu_1(A)+\mu_1(\{x\}) \leqq \sup_{B\subset W\setminus\{x\}} \mu_1(B)+\mu_1(\{x\}).$$

Hence by (5) we obtain

$$\mu_1(\bar{H}) < 1+ \sup_{\mu\in\mathcal{M}} \mu(\{x\}). \tag{6}$$

Let $E_1=U\setminus\bar{H}$. Then we have $E_1\subset O$ and $E_1\cap\bar{H}=\emptyset$. By the inequality

$$\mu_1(E_1)+\mu_1(\bar{H}) \geqq \mu_1(U),$$

(4) and (6) we obtain $\mu_1(E_1)>1$. Using the preceding procedure, taking in the inequality (3) the constant 5 instead of 4 and taking into account the facts that $x\in H$ and the family $\mathcal{M}$ is not bounded on H, we obtain open sets E_2, H_1 ($H_1\subset H$) and $\mu_2\in\mathcal{M}$ such that $E_2\cap H_1=\emptyset$, $x\in H_1$ and $\mu_2(E_2)>2$. We have $E_1\cap E_2=\emptyset$. Continuing this procedure we obtain a sequence (μ_i) from $\mathcal{M}$ and a sequence (E_i)

of pairwise disjoint open sets such that

(7) $$\mu_i(E_i) > i \quad (i \in N).$$

We shall prove that μ_i $(i \in N)$ are exhaustive on the sequence (E_n) of disjoint open sets, i.e.

(8) $$\lim_{j \to \infty} \mu_i(E_j) = 0 \quad (i \in N).$$

Since $\bigcup_{j=1}^{\infty} E_j$ is an open set and $|\mu_i|$ are regular, for $\varepsilon > 0$ by Corollary 1 there exists a compact set $K' \subset \bigcup_{j=1}^{\infty} E_j$ such that $\mu_i(C) < \varepsilon$ for each $i \in N$ and each $C \subset \bigcup_{j=1}^{\infty} E_j \setminus K'$. Since (E_i) is an open cover of K', there exists $n_0 \in N$ such that $K' \subset \bigcup_{j=1}^{n_0} E_j$. Then we have for $m \geqq n_0 + 1$

$$\mu_i(E_m) \leqq \sup_{C'} \mu_i(C') \leqq \sup_{C} \mu_i(C) < \varepsilon \quad (i \in N)$$

where $C' \subset E_m \cup \left(\bigcup_{j=1}^{n_0} E_j \setminus K'\right)$ and $C \subset \bigcup_{j=1}^{\infty} E_j \setminus K'$. So we obtain (8).

Let $x_{ij} = \mu_i(E_j)/i$. We have by (8) $\lim_{j \to \infty} x_{ij} = 0$ $(i \in N)$. We obtain by the boundedness assumption of the theorem that $\lim_{i \to \infty} x_{ij} = 0$ $(j \in N)$. Applying Lemma 2 for the infinite matrix $[x_{ij}]$ $(i, j \in N)$ we obtain a sequence (i_n) from N and a sequence (ε_n) of positive real numbers such that

(9) $$\sum_{k=1}^{n-1} x_{i_n i_k} = (1/2 - \varepsilon_n) x_{i_n i_n},$$

(10) $$x_{i_n i_{n+q}} < 2^{-q} \varepsilon_n x_{i_n i_n} \quad (n \in N).$$

Using the triangularity of μ_{i_n} $(n \in N)$ and Lemma 1 we obtain

$$\mu_{i_n}\left(\bigcup_{k=1}^{\infty} E_{i_k}\right) \geqq \mu_{i_n}(E_{i_n}) - \sum_{k=1}^{n-1} \mu_{i_n}(E_{i_k}) - \sum_{k=n+1}^{\infty} \mu_{i_n}(E_{i_k}) \quad (n \in N).$$

Hence by (9) and (10)

$$i_n^{-1} \mu_{i_n}\left(\bigcup_{k=1}^{\infty} E_{i_k}\right) \geqq x_{i_n i_n} - \sum_{k=1}^{n-1} x_{i_n i_k} - \sum_{k=n+1}^{\infty} x_{i_n i_k} \geqq x_{i_n i_n}/2 \quad (n \in N),$$

i.e.,

$$\mu_{i_n}\left(\bigcup_{k=1}^{\infty} E_{i_k}\right) \geqq \mu_{i_n}(E_{i_n})/2 \quad (n \in N).$$

Then by (7) we obtain

$$\mu_{i_n}\left(\bigcup_{k=1}^{\infty} E_{i_k}\right) \geqq i_n/2 \quad \text{for each } n \in N.$$

Since $\bigcup_{k=1}^{\infty} E_{i_k}$ is an open set we obtain a contradiction with the boundedness of (μ_{i_n}) on open sets.

Finally, we reduce the general case $k\in(0, +\infty)$ to the preceding one. Namely, for $r\in(0, 1]$ this follows from the fact that each r-triangular set function is 1-triangular. Let us suppose now that $\mu_n{}'$ $(n\in\mathrm{N})$ are k-triangular for some k such that $k>1$. Since for any number k such that $k>1$ and any $r\in(0, 1]$ there exists $m\in\mathrm{N}$ such that $k\leqq mr$, it follows that the set functions v_n, $v_n=m\cdot\mu_n$ $(n\in\mathrm{N})$, are r-triangular. Now an application of the first part to the set functions $\{v_n\}$ completes the proof.

4. Semigroup valued k-triangular set functions

Let X be a commutative semigroup with a neutral element O. Let d: $X\to[0, +\infty)$ be a pseudometric such that satisfies the following condition

$$(\mathrm{d}_+) \qquad d(x+x_1, y+y_1) \leqq d(x, y)+d(x_1, y_1)$$

for all $x, x_1, y, y_1\in X$.

Example. WEBER [13] has proved that for every commutative complete uniform semigroup there exists a family of pseudometrics which satisfy (d_+) and which generate its uniformity.

Let X be endowed with a pseudometric d which satisfies (d_+). Now we can extend the definition of the regularity of a set function v: $\mathscr{S}\to X$ taking only in the Definition 1 v and "$d(v(A), v(A'))<\varepsilon$" instead of μ and "$|v(A)-v(A')|<\varepsilon$", respectively.

The pseudometric d induces a *triangular functional* [8], [10] in the following way:

$$f(x):=d(x, 0) \quad (x\in X).$$

The functional f satisfies

(F_1) $f(x+y)\leqq f(x)+f(y)$, and

(F_2) $f(x+y)\geqq |f(x)-f(y)|$ for all $x, y\in X$.

Now we define the variation $|v|$ of a set function v: $\mathscr{S}\to X$ with $v(\emptyset)=0$ in the following way:

$$|v|(E):=\sup_{\pi}\sum_{A\in\pi} f(v(A)) \quad (E\in\mathscr{S})$$

where the supremum is taken over all partitions π of E into a finite number of pairwise disjoint members of $\mathcal{S}$. It is easy to see that $|\nu|$ is superadditive.

A set function $\nu: \mathcal{S} \to X$ is said to be a semigroup valued k-triangular set function if satisfies $\nu(\emptyset)=0$, and

$$f(\nu(A)) - kf(\nu(B)) \leqq f(\nu(A \cup B)) \leqq f(\nu(A)) + kf(\nu(B))$$

for $A, B \in \mathcal{S}$ with $A \cap B = \emptyset$.

Now we have the following generalization of Theorem 2.

Theorem 3. *Let $\mathcal{F}$ be a family of semigroup valued k-triangular set functions with regular variations defined on $\mathcal{B}$. If the set $\{f(\nu(O));\ \nu \in \mathcal{F}\}$ is bounded for every open set O, then*

$$\{f(\nu(B));\ \nu \in \mathcal{F},\ B \in \mathcal{B}\}$$

is a bounded set.

Proof. We take $\mu(B) := f(\nu(B))$ $(B \in \mathcal{B}, \nu \in \mathcal{F})$ and we apply Theorem 2.

Remark 2. Diagonal theorems ([1], [8], [9], [12]) are very useful in the elementary proofs of many important theorems in functional analysis and measure theory.

References

[1] P. Antosik, On the Mikusiński diagonal theorem, *Bull. Acad. Polon. Sci., Ser. Math. Astronom. Phys.*, **19** (1971), 305—310.

[2] J. Diestel, J. J. Uhl, *Vector measures*, Math. Surveys 15, Amer. Math. Soc. (Providence, R. I., 1977).

[3] J. Dieudonné, Sur la convergence des suites de mesures de Radon, *Anais Acad. Brasil. Ci.*, **23** (1951), 21—38, 277—281.

[4] N. Dinculeanu, *Vector Measures*, Pergamon Press (New York, 1967).

[5] L. Drewnowski, On the continuity of certain non-additive set functions, *Coll. Math.*, **38** (1978), 243—253.

[6] Н. С. Гусельников, О продолжении квазилипшицвых функций множества, *Математические заметки*, **17** (1975), 21—31.

[7] J. Kupka, Uniform boundedness principles for regular Borel vector measures, *J. Austral. Math. Soc. Series A*, **29** (1980), 206—218.

[8] E. Pap, A generalization of the diagonal theorem on a block-matrix, *Mat. ves.*, **11** (26) (1974), 66—71.

[9] E. Pap, On the continuous functions with the values in a semigroup, in: *Topology* (Proc. Conf. Budapest, 1978), Coll. Math. Soc. János Bolyai, Vol. 23, North-Holland (Amsterdam, 1980); pp. 931—938.

[10] E. Pap, Triangular functionals on commutative semigroups, in: *Proc. of the Conf. on Convex Structures in Schwerin*, Akad. der Wiss. der DDR (to appear).

[11] J. D. STEIN, A uniform-boundedness theorem for measures, *Mich. Math. J.*, **19** (1972), 161—165.

[12] C. SWARTZ, Applications of the Mikusiński's diagonal theorem, *Bull. Acad. Polon. Sci. Ser. Math. Astronom. Phys.*, **26** (1978), 421—424.

[13] W. WEBER, Fortsetzung von Massen mit Werten in uniformen Halbgruppen, *Arch. Math.*, **27** (1976), 412—423.

INSTITUTE OF MATHEMATICS
ILIJE DJURIČIĆA 4
21 000 NOVI SAD, YUGOSLAVIA

Acta Sci. Math., 50 (1986), 169—182

On the asymptotic estimate of the maximum likelihood of parameters of the spectral density having zeros

M. S. GINOVIAN

1. Introduction

1. Let x_t, $t=0, \pm 1, \cdots$ be a stationary Gaussian time-series with $E(x_t)=0$ and spectral density (SD) $f(\lambda)$. Suppose that the SD f is a function of an unknown vector parameter $\theta=(\theta_1, \ldots, \theta_p)' \in \Theta$, where Θ is a bounded closed set in the p-dimensional Euclidean space R^p. We wish to obtain an estimate of this parameter from data consisting of a part of a realisation of the series, which will be assumed to be n consecutive observations denoted by $x_1, \ldots, x_n$. Obviously, we can consider the maximum likelihood estimate $\tilde{\theta}_n$ of the parameter θ:

(1) $$L_n(\tilde{\theta}_n; X) = \max_{\theta} L_n(\theta; X),$$

where $L_n(\theta; X)$ is the logarithm of the likelihood of the data $X=(x_1, \ldots, x_n)'$. The function $L_n(\theta; X)$ can be written in the form (see [9], [11])

(2) $$L_n(\theta; X) = -(1/2)\{n \ln 2\pi + \ln \det B_{n,f_\theta} + X' B_{n,f_\theta}^{-1} X\},$$

where $B_{n,f_\theta} = \|c_{k-j}(\theta)\|_{k,j=\overline{1,n}}$ is the Toeplitz matrix connected with the function $f(\lambda; \theta)$.

It follows from formulas (1) and (2) that in order to find the estimate $\tilde{\theta}_n$, it is necessary to obtain the explicit expressions for $\det B_{n,f_\theta}$ and B_{n,f_θ}^{-1}, and this is a very difficult problem. Even in the simplest case of the first order autoregression the explicit expression for $L_n(\theta; X)$ is complicated (see [3]).

Following WHITTLE [12] and WALKER [11], let us introduce the estimate $\hat{\theta}_n$ of parameter θ:

(3) $$\tilde{L}_n(\hat{\theta}_n; X) = \max_{\theta} \tilde{L}_n(\theta; X),$$

Received January 11, 1984.

where $\tilde{L}_n(\theta; X)$ is "the main part" of the function $L_n(\theta; X)$ satisfying the condition

(4) $$n^{-1/2}[L_n(\theta; X)-\tilde{L}_n(\theta; X)] \to 0 \quad \text{as} \quad n \to \infty,$$

where the convergence is in probability. The estimate $\hat{\theta}_n$ will be called the asymptotical estimate of the maximum likelihood (AEML).

For the strictly positive SD case the asymptotical properties of the AEML $\hat{\theta}_n$ were investigated by WALKER [11] and DZHAPARIDZE [3]. The case in which the SD has "weak" zeros independent of the parameter θ was considered by the author [5]. In these papers it was shown that under wide conditions on the SD the function $\tilde{L}_n(\theta; X)$ can be chosen to have a much simpler form than $L_n(\theta; X)$. Moreover the estimates $\tilde{\theta}_n$ and $\hat{\theta}_n$ are asymptotically equivalent, i.e. the estimate $\hat{\theta}_n$ is also consistent, asymptotically normal and asymptotically efficient.

In the present paper we generalize the abovementioned results to the case in which the SD has both "weak" and "strong" zeros of polynomial type, i.e. when the function $f(\lambda; \theta)$ admits the representation

(5) $$f(\lambda; \theta) = |Q_m(e^{i\lambda})|^2 h(\lambda; \theta),$$

where $Q_m(e^{i\lambda})$ $(|Q_m(0)|=1)$ is a polynomial of degree m with roots on the unit circle, which are independent of the parameter θ, and the function $h(\lambda; \theta)$ has "weak" zeros also independent of the parameter θ.

Note that a similar case was considered by DZHAPARIDZE [3], [4], but under stronger restrictions on the function $f(\lambda; \theta)$. Namely he assumed that the function $h(\lambda; \theta)$ is strongly positive and the polynomial $Q_m(e^{i\lambda})$ has no multiple roots.

2. The following notations will be used: L_f^2 is the weight L^2 space with weight f; $H_n(f)$ is the space of polynomials of degree n, considered as a subspace of L_f^2; P_n^f is the projector from L_f^2 to $H_n(f)$; $G_n^f(\lambda, \mu)$ is the reproducing kernel of the space $H_n(f)$; $\|\cdot\|_f$ and $(\cdot, \cdot)_f$ are respectively the norm and inner product in L_f^2; $|A|_f$ and $\|A\|_f$ are respectively the uniform and Hilbert—Schmidt norms of the operator A in L_f^2.

Remark. In all notations the symbol f will be omitted if $f(\lambda)\equiv 1$.

We shall use the main result of [5], therefore for completeness of presentation we reproduce it here.

Theorem A [5]. *Let the* SD $f(\lambda; \theta)$ *of the stationary Gaussian time-series* x_t *admit the representation* (5), *where* $Q_m(e^{i\lambda})$ $(|Q_m(0)|=1)$ *is a polynomial of degree* m *with roots on the unit circle, which are independent of* θ, *and the function* $h(\lambda; \theta)$ *satisfies the following conditions:*

1. $\ln h(\lambda; \theta) = u(\lambda; \theta)+\tilde{v}(\lambda; \theta), \quad \theta \in \Theta,$

where $u(\cdot\,;\theta)$ *and* $v(\cdot\,;\theta)$ *are bounded functions* ($\tilde{v}$ *is the harmonic conjugate of* v) *and* $\|v\|_\infty<\pi/2$;

2. $\sum_{|k|>n}|a_k(\theta)|^2=o\big(1/\sqrt{n}\big),\quad n\to\infty;$

3. $\sum_{|k|>n}|c_k(\theta)|^2=o\big(1/(\sqrt{n}\ln n)\big),\quad n\to\infty;$

4. $\sum_{|k|>n}|b_k(\theta)|^2=o\big(1/(\sqrt{n}\ln n)\big),\quad n\to\infty,$

for all $\theta\in\Theta$, *where* $a_k(\theta)$, $c_k(\theta)$ *and* $b_k(\theta)$ *are the Fourier coefficients of the functions* $\ln h(\cdot\,;\theta)$, $h(\cdot\,;\theta)$ *and* $1/h(\cdot\,;\theta)$, *respectively. Then the limiting relation* (4) *holds, where the function* $L_n(\theta;X)$ *is given by* (2) *and*

(6)
$$\tilde{L}_n(\theta;X)=-(n/2)\Big\{\ln 2\pi+(1/2\pi)\int_{-\pi}^{\pi}\ln h(\lambda;\theta)\,d\lambda+(1/2\pi)\int_{-\pi}^{\pi}\big(\tilde{I}_n(\lambda)/h(\lambda;\theta)\big)\,d\lambda\Big\},$$

where

(7)
$$\tilde{I}_n(t)=(1/n)\iint_{-\pi}^{\pi}G_n^{|Q_m|^2}(\lambda,t)\,G_n^{|Q_m|^2}(t,\mu)\,|Q_m(t)|^2Z^f(d\lambda)\overline{Z^f(d\mu)}$$

is the generalized periodogram of x_t. ($Z^f(d\lambda)$ *is the orthogonal stochastic measure participating in the spectral representation of* x_t: $x_t=\int_{-\pi}^{\pi}\exp(i\lambda t)Z^f(d\lambda)$.)

2. Auxiliary results

Let the functions $f(\lambda;\theta)$ and $h(\lambda;\theta)$ be connected by the relation (5). We have the obvious inclusion

$$Q_mH_n(f)\subseteqq H_{n+m}(h).$$

Let us denote by N_m the orthogonal complement of $Q_mH_n(f)$ in $H_{n+m}(h)$:

(8)
$$H_{n+m}(h)=Q_mH_n(f)\oplus N_m.$$

Denoting by $G_{n+m}(\lambda,\mu)$, $G_n^{|Q_m|^2}(\lambda,\mu)$ and $R_m(\lambda,\mu)$ the reproducing kernels of the spaces H_{n+m}, $H_n(|Q_m|^2)$ and N_m, respectively, from (8) we have

(9)
$$G_{n+m}(\lambda;\mu)=Q_m(\lambda)\overline{Q_m(\mu)}\,G_n^{|Q_m|^2}(\lambda,\mu)+R_m(\lambda,\mu).$$

From (8) we also obtain that

(10)
$$P_{n+m}=\Phi_n+T_m,$$

where P_{n+m}, Φ_n and T_m are the projectors from L^2 to subspaces H_{n+m}, $Q_m H_n(|Q_m|^2)$ and N_m, respectively.

The following assumptions will be made throughout the paper.

A1. The true value θ_0 of the parameter θ belongs to a bounded closed set Θ contained in an open set S in the p-dimensional Euclidean space R^p.

A2. If θ_1 and θ_2 are any two points of Θ, $f(\lambda; \theta_1)$ and $f(\lambda; \theta_2)$ are not equal almost everywhere (λ).

A3. For SD$f(\lambda; \theta)$ all the conditions of Theorem A are satisfied.

Lemma 1. *Let the partial derivatives* $\partial \ln h(\lambda; \theta)/\partial\theta_k$, $k=\overline{1,p}$, *be continuous functions of* (λ, θ) *for* $\lambda\in[-\pi, \pi]$, $\theta\in S$. *Then for any* $\theta_1\in\Theta$ *such that* $\theta_1\neq\theta_0$ *(*θ_0 *is the true value of* θ*)*

(11)

$$\lim_{n\to\infty} (1/n) \iint_{-\pi}^{\pi} |G_n^{|Q_m|^2}(\lambda, t) Q_m(\lambda)\overline{Q_m(t)}|^2 r(t; \theta_0, \theta_1)(h(\lambda; \theta_0)/h(t; \theta_0))\, d\lambda\, dt = \int_{-\pi}^{\pi} r(t; \theta_0, \theta_1)\, dt,$$

where

$$r(t; \theta_0, \theta_1) = 1 - h(t; \theta_0)/h(t; \theta_1).$$

Proof. Under the given conditions we have (see [5, proof of Lemma 6])

$$\lim_{n\to\infty} (1/n) \iint_{-\pi}^{\pi} |G_{n+m}(\lambda, t)|^2 r(t; \theta_0, \theta_1)(h(\lambda; \theta_0)/h(t; \theta_1))\, d\lambda\, dt = \int_{-\pi}^{\pi} r(t; \theta_0, \theta_1)\, dt. \tag{12}$$

To prove Lemma 1 it therefore suffices to show that

$$\lim_{n\to\infty} (1/n) \iint_{-\pi}^{\pi} [|G_{n+m}(\lambda, t)|^2 - |G_n^{|Q_m|^2}(\lambda, t) Q_m(\lambda)\overline{Q_m(t)}|^2]\times r(t; \theta_0, \theta_1)(h(\lambda; \theta_0)/h(t; \theta_1))\, d\lambda\, dt = 0. \tag{13}$$

It is easy to see that

$$\iint_{-\pi}^{\pi} [|G_{n+m}(\lambda, t)|^2 - |G_n^{|Q_m|^2}(\lambda, t) Q_m(\lambda)\overline{Q_m(t)}|^2]\times r(t; \theta_0, \theta_1)(h(\lambda; \theta_0)/h(t; \theta_1))\, d\lambda\, dt = \operatorname{tr}(P_{n+m}(r_{01}/h_0)P_{n+m}h_0 - \Phi_n(r_{01}/h_0)\Phi_n h_0), \tag{14}$$

where r_{01}/h_0 and h_0 are the operators of multiplication by the functions $r(t; \theta_0, \theta_1)/h(t; \theta_0)$ and $h(t; \theta_0)$, respectively.

Therefore using formula (10) we obtain

$$\operatorname{tr}\left(P_{n+m}(r_{01}/h_0)P_{n+m}h_0-\Phi_n(r_{01}/h_0)\Phi_n h_0\right)= \tag{15}$$
$$=\operatorname{tr}\left(\Phi_n(r_{01}/h_0)T_m h_0+T_m(r_{01}/h_0)\Phi_n h_0+T_m(r_{01}/h_0)T_m h_0\right).$$

Further, using the inequalities (see e.g. [2])

$$\operatorname{tr}(A_n B_n)\leqq \|A_n\|_h\|B_n\|_h, \tag{16}$$

$$\|A_n\|_h\leqq \sqrt{n}\,|A_n|_h \tag{17}$$

and the relation $|A_n|_{1/h}=|(1/h)A_n h|_h$, from (15) we find

$$(1/n)\left|\operatorname{tr}\left(P_{n+m}(r_{01}/h_0)P_{n+m}h_0-\Phi_n(r_{01}/h_0)\Phi_n h_0\right)\right|\leqq \tag{18}$$
$$\leqq \left(\sqrt{m/n}\sup_{\theta,\lambda}|r(\lambda;\,\theta)|\left[|\Phi_n|_h|T_m|_{1/h}+|\Phi_n|_{1/h}|T_m|_h+\sqrt{m/n}\,|T_m|_{1/h}|T_m|_h\right]\right.$$

The right-hand side of (18) tends to zero as $n\to\infty$, since by Lemma 1 in [5] and by formula (10)

$$\sup_n|\Phi_n|_h<\infty \quad\text{and}\quad \sup_n|\Phi_n|_{1/h}<\infty.$$

Lemma 2. *Under the conditions of Lemma* 1

$$\lim_{n\to\infty}(1/n)\int_{-\pi}^{\pi}\!\!\int\Big|\int_{-\pi}^{\pi}G_n^{|Q_m|^2}(\lambda,t)G_n^{|Q_m|^2}(t,\mu)Q_m(\lambda)\overline{Q_m(\mu)}\times \tag{19}$$
$$\times|Q_m(t)|^2\left(r(t;\,\theta_0,\theta_1)/h(t;\,\theta_0)\right)dt\Big|^2 h(\lambda;\,\theta_0)h(\mu;\,\theta_0)\,d\lambda\,d\mu=\int_{-\pi}^{\pi}r^2(t;\,\theta_0,\theta_1)\,dt.$$

Proof. It is known that under the given conditions (see [5, proof of Lemma 6])

$$\lim_{n\to\infty}(1/n)\int_{-\pi}^{\pi}\!\!\int\Big|\int_{-\pi}^{\pi}G_{n+m}(\lambda,t)G_{n+m}(t;\,\mu)\left(r(t;\,\theta_0,\theta_1)/h(t;\,\theta_0)\right)dt\Big|^2\times$$
$$\times h(\lambda;\,\theta_0)h(\mu;\,\theta_0)\,d\lambda\,d\mu=\int_{-\pi}^{\pi}r^2(t;\,\theta_0,\theta_1)\,dt.$$

Hence to prove Lemma 2 it is enough to show that

$$\lim_{n\to\infty}(1/n)\int_{-\pi}^{\pi}\!\!\int\Big[\Big|\int_{-\pi}^{\pi}G_{n+m}(\lambda,t)G_{n+m}(t,\mu)\left(r(t;\,\theta_0,\theta_1)/h(t;\,\theta_0)\right)dt\Big|^2- \tag{20}$$
$$-\Big|\int_{-\pi}^{\pi}G_n^{|Q_m|^2}(\lambda,t)G_n^{|Q_m|^2}(t,\mu)Q_m(\lambda)\overline{Q_m(\mu)}|Q_m(t)|^2\times$$
$$\times\left(r(t;\,\theta_0,\theta_1)/h(t;\,\theta_0)\right)dt\Big|^2\Big]h(\lambda;\,\theta_0)h(\mu;\,\theta_0)\,d\lambda\,d\mu=0.$$

It is easy to see that the relation (20) is equivalent to the following

$$\lim_{n\to\infty} [\|P_{n+m}(r_{01}/h_0)P_{n+m}h_0\|_h^2 - \|\Phi_n(r_{01}/h_0)\Phi_n h_0\|_h^2] = 0. \tag{21}$$

Using the well-known inequality (see [1])

$$\|A_n\|_h^2 - \|B_n\|_h^2 \leqq \|A_n - B_n\|_h^2 + 2\|A_n - B_n\|_h \|B_n\|_h$$

and the fact that

$$\|\Phi_n(r_{01}/h_0)\Phi_n h_0\|_h \leqq \sqrt{n} \sup_{\lambda,\theta} |r(\lambda;\, \theta)|\, |\Phi_n|_h |\Phi_n|_{1/h} = o(\sqrt{n})$$

when $n\to\infty$ (which follows from inequality (17) and Lemma 1 in [5]), it is easy to see that to prove (21) it suffices to show that

$$\lim_{n\to\infty} (1/n)\|P_{n+m}(r_{01}/h_0)P_{n+m}h_0 - \Phi_n(r_{01}/h_0)\Phi_n h_0\|_h^2 = 0. \tag{22}$$

From formula (10) we have

$$P_{n+m}(r_{01}/h_0)P_{n+m}h_0 - \Phi_n(r_{01}/h_0)\Phi_n h_0 = T_m(r_{01}/h_0)P_{n+m}h_0 + \Phi_n(r_{01}/h_0)T_m h_0.$$

Using this fact, the inequalities (16) and $\|A_nB_n\|_h \leqq \|A_n\|_h |B_n|_h$, we find

$$\begin{aligned} &(1/n)\|P_{n+m}(r_{01}/h_0)P_{n+m}h_0 - \Phi_n(r_{01}/h_0)\Phi_n h_0\|_h^2 \leqq \\ &\leqq (2/n)\{\|T_m\|_h^2 |(r_{01}/h_0)P_{n+m}h_0|_h^2 + |\Phi_n|_h^2 \|(r_{01}/h_0)T_m h_0\|_h^2\} \leqq \\ &\leqq (2/n)\sup_{\lambda,\theta}|r(\lambda;\,\theta)|\Big\{\|T_m\|_h^2 \sup_n |P_{n+m}|_{1/h}^2 + \|T_m\|_{1/h}^2 \sup_n |\Phi_n|_h^2\Big\} \to 0 \end{aligned}$$

as $n\to\infty$, since by Lemma 1 in [5] and formula (10)

$$\sup_n |\Phi_n|_h < \infty \quad \text{and} \quad \sup_n |P_{n+m}|_{1/h} < \infty.$$

Lemma 3. *Let A_n be an $n\times n$ matrix such that $|A_n|\to 0$ as $n\to\infty$ and $\sup_n \|A_n\| < \infty$. Then*

$$\lim_{n\to\infty} [\ln (E_n + A_n) - \operatorname{tr}(A_n) + (1/2)\|A_n\|^2] = 0.$$

The proof easily follows from inequality (V) in [2].

3. The asymptotic properties of AEML $\hat{\theta}_n$

It follows from Theorem A, that the AEML $\hat{\theta}_n$ can be found from the relation

$$U_n(\hat{\theta}_n;\, X) = \min_\theta U_n(\theta;\, X),$$

where

$$U_n(\theta;\, X) = (1/4\pi) \int_{-\pi}^{\pi} [\ln h(\lambda;\, \theta) + (\tilde{I}_n(\lambda)/h(\lambda;\, \theta))]\, d\lambda. \tag{23}$$

1. Consistency of AEML $\hat{\theta}_n$.

Theorem 1. *Let the partial derivatives* $\partial \ln h(\lambda; \theta)/\partial\theta_k$, $k=\overline{1,p}$, *be continuous functions of* (λ, θ) *for* $\lambda\in[-\pi, \pi]$, $\theta\in S$. *Then the* AEML $\hat{\theta}_n$ *is consistent, i.e.* $\hat{\theta}_n\to\theta_0$, *as* $n\to\infty$, *in probability.*

We first establish three lemmas.

Lemma 4. *Let* θ_0 *be the true value of* θ *and* θ_1 *be any other point of* Θ. *Then there is a positive constant* $K(\theta_0, \theta_1)$ *such that*

$$\lim_{n\to\infty} \mathrm{P}\{U_n(\theta_0)-U_n(\theta_1) < -K(\theta_0, \theta_1)\} = 1. \tag{24}$$

Proof. From formulas (7) and (23) we have

$$W_n \stackrel{\text{def}}{=} U_n(\theta_0)-U_n(\theta_1) =(1/4\pi) \int_{-\pi}^{\pi} (\ln h(t;\ \theta_0)/h(t;\ \theta_1))\,dt+ \tag{25}$$

$$+(1/4\pi) \int_{-\pi}^{\pi} \tilde{I}_n(t)[1/h(t;\ \theta_0)-1/h(t;\ \theta_1)]\,dt =$$

$$=(1/4\pi) \int_{-\pi}^{\pi} (\ln h(t;\ \theta_0)/h(t;\ \theta_1)\,dt+(1/4\pi n)\iiint_{-\pi}^{\pi} G_n^{|Q_m|^2}(\lambda, t)\,G_n^{|Q_m|^2}(t, \mu)\times$$

$$\times|Q_m(t)|^2[1/h(t;\ \theta_0)-1/h(t;\ \theta_1)]Z^f(d\lambda)\overline{Z^f(d\mu)}.$$

Hence

$$\mathrm{E}(W_n) = (1/4\pi) \int_{-\pi}^{\pi} (\ln h(t;\ \theta_0)/h(t;\ \theta_1))\,dt+$$

$$+(1/4\pi n) \iint_{-\pi}^{\pi} |G_n^{|Q_m|^2}(\lambda, t)|^2\,|Q_m(t)|^2[1/h(t;\ \theta_0)-1/h(t;\ \theta_1)]\,f(\lambda;\ \theta_0)\,d\lambda\,dt.$$

Now using Lemma 1 we obtain

(26)
$$\mathrm{E}(W_n) = (1/4\pi) \int_{-\pi}^{\pi} \ln(h(t;\ \theta_0)/h(t;\ \theta_1))\,dt+(1/4\pi) \int_{-\pi}^{\pi} [1-h(t;\ \theta_0)/h(t;\ \theta_1)]\,dt+o(1).$$

By the obvious inequality

$$\ln(h(t;\theta_0)/h(t;\theta_1)) < (h(t;\theta_0)/h(t;\theta_1))-1 \tag{27}$$

(here by assumption A2 we have strict inequality) from (26) we obtain

$$\lim_{n\to\infty} \mathrm{E}(W_n) \stackrel{\text{def}}{=} -l(\theta_0, \theta_1), \quad \text{say, where} \quad l(\theta_0, \theta_1) > 0.$$

Also, from (25) we find

$$\mathrm{D}(\sqrt{n}\,W_n) = (1/4\pi n) \iint_{-\pi}^{\pi} \Big| \int_{-\pi}^{\pi} G_n^{|Q_m|^2}(\lambda, t) G_n^{|Q_m|^2}(t, \mu) \times$$

$$\times |Q_m(t)|^2 [1/h(t;\, \theta_0) - 1/h(t;\, \theta_1)]\, dt \Big|^2 f(\lambda;\, \theta_0) f(\mu;\, \theta_0)\, d\lambda\, d\mu.$$

Hence by Lemma 2 we get

$$\lim_{n\to\infty} \mathrm{D}(\sqrt{n} W_n) = (1/4\pi) \int_{-\pi}^{\pi} [1 - h(t;\, \theta_0)/h(t;\, \theta_1)]^2\, dt < \infty.$$

The desired result (24) then follows by a simple application of Chebyshev's inequality: $K(\theta_0, \theta_1)$ can be any constant less than $l(\theta_0, \theta_1)$.

Lemma 5. *Let $\theta_1 \in \Theta$ and $\theta_2 \in S$ be chosen such that $|\theta_2 - \theta_1| < \delta$ (δ possibly depending on θ_1). Then there exists a number $n_0 > 0$ such that for $n \geqq n_0$*

$$|U_n(\theta_1; X) - U_n(\theta_2; X)| \leqq H_{\delta, n}(\theta_1; X), \tag{28}$$

where $H_{\delta,n} = H_{\delta,n}(\theta_1; X)$ is a random variable such that

$$\lim_{\delta\to 0} \mathrm{E}(H_{\delta,n}) = 0 \quad \textit{uniformly in} \quad n \geqq n_0, \tag{29}$$

and

$$\lim_{n\to\infty} \mathrm{D}(H_{\delta,n}) = 0. \tag{30}$$

Proof. From (25), using inequality (27), we obtain

$$|U_n(\theta_1) - U_n(\theta_2)| \leqq (1/4\pi) \int_{-\pi}^{\pi} [1 + \tilde{I}_n(\lambda)/h(\lambda;\, \theta_1)] \big| \ln \big(h(\lambda;\, \theta_1)/h(\lambda;\, \theta_2)\big) \big|\, d\lambda. \tag{31}$$

Let us denote by

$$H_{\theta_1, \delta(\theta_1)} = \sum_{k=1}^{p} \sup_{-\pi \leqq \lambda \leqq \pi} \sup_{|\theta_1 - \theta| < \delta(\theta_1)} |\partial \ln h(\lambda;\, \theta)/\partial \theta_k|,$$

where $\delta(\theta_1) > \delta$ is chosen so that the set $\{\theta;\, |\theta_1 - \theta| \leqq \delta(\theta_1)\}$ is contained in S. Then, by the mean value theorem, from (31) we get

$$|U_n(\theta_1) - U_n(\theta_2)| \leqq H_{\delta,n}(\theta_1; X),$$

where

$$H_{\delta,n}(\theta_1;\, X) = (\delta/4\pi) H_{\theta_1, \delta(\theta_1)} \int_{-\pi}^{\pi} [1 + \tilde{I}_n(t)/h(t;\, \theta_1)]\, dt. \tag{32}$$

We now show that the random variable $H_{\delta,n}(\theta_1; X)$ satisfies the conditions (29) and (30). From (32) we have

$$\mathrm{E}(H_{\delta,n}) = (\delta/2) H_{\theta_1, \delta(\theta_1)} + (\delta/4\pi n) H_{\theta_1, \delta(\theta_1)} \times \tag{33}$$

$$\times \iint_{-\pi}^{\pi} |G_n^{|Q_m|^2}(\lambda, t) Q_m(t)|^2 \big(f(\lambda;\, \theta_1)/h(t;\, \theta_1)\big)\, d\lambda\, dt$$

and

(34) $$D(H_{\delta,n}) = ((\delta/4\pi n)H_{\theta_1,\delta(\theta_1)})^2 \iint_{-\pi}^{\pi} \Big| \int_{-\pi}^{\pi} (G_n^{|Q_m|^2}(\lambda;\, t) G_n^{|Q_m|^2}(t,\mu)\times$$
$$\times |Q_m(t)|^2/h(t;\, \theta_1))\, dt\Big|^2 f(\lambda;\, \theta_1) f(\mu;\, \theta_1)\, d\lambda\, d\mu.$$

So it is easy to see that (29) follows from (33) and Lemma 1 while (30) follows from (34) and Lemma 2.

Lemma 6 (WALKER [11]). *Let the random variable $U_n(\theta)$ satisfy the relation* (28) *for all $\theta_1 \in \Theta$, $\theta_2 \in S$ such that $|\theta_1-\theta_2|<\delta$ (δ possibly depending on θ_1) and $H_{\delta,n}(\theta_1; X)$ satisfies the relations* (29) *and* (30). *Then*

$$\mathrm{P}\lim_{n\to\infty} \hat\theta_n = \theta_0, \quad \textit{the true value of } \theta.$$

The proof of Theorem 1 now immediately follows from Lemmas 4, 5 and 6.

2. Asymptotic normality and asymptotic efficiency of AEML $\hat\theta_n$. Having established the consistency of AEML $\hat\theta_n$, we can go on to obtain the limiting distribution of the vector $\sqrt{n}\,(\hat\theta_n-\theta_0)$ in the usual way by applying the mean value theorem to $U_n^{(i)}(\hat\theta_n)-U_n^{(i)}(\theta_0)$, $i=\overline{1,p}$, where $U_n^{(i)}$ denotes the partial derivative $\partial U_n(\theta)/\partial\theta_i$, $(U_n^{(i)}(\theta_0)=\partial U_n(\theta)/\partial\theta_i|_{\theta=\theta_0})$, and θ_0 is the true value of θ. Of course, further conditions must be imposed on SD to ensure that the second order partial derivatives $U_n^{(ij)}=\partial^2 U_n/\partial\theta_i\partial\theta_j$ satisfy a suitable continuity condition and that a central limit theorem can be applied to give the limiting joint distribution of $\sqrt{n}\, U_n^{(i)}(\theta_0)$, $i=\overline{1,p}$.

Theorem 2. *Let the functions $\partial \ln h(\lambda;\theta)/\partial\theta_k$, $k=\overline{1,p}$, be continuous in (λ,θ) for $\lambda\in[-\pi,\pi]$, $\theta\in S$, and the functions $\partial^2 \ln h(\lambda;\theta)/\partial\theta_j\partial\theta_k$, $\partial^3 \ln h(\lambda;\theta)/\partial\theta_l\partial\theta_j\partial\theta_k$, $k,j,l=\overline{1,p}$, be continuous in (λ,θ) for $\lambda\in[-\pi,\pi]$, $\theta\in N_\delta(\theta_0)$, where $N_\delta(\theta_0)=\{\theta;\, |\theta-\theta_0|<\delta\}$ is some neighbourhood of θ_0, and let the matrix $\Gamma_0=\|\gamma_{ij}(\theta_0)\|_{i,j=\overline{1,p}}$ with*

$$\gamma_{ij}(\theta_0) = (1/4\pi) \int_{-\pi}^{\pi} (\partial \ln h(\lambda;\, \theta)/\partial\theta_i)_{\theta=\theta_0} (\partial \ln h(\lambda;\, \theta)/\partial\theta_j)_{\theta=\theta_0}\, d\lambda$$

be non-singular. Then the limiting distribution of the vector $\sqrt{n}\,(\hat\theta_n-\theta_0)$ when $n\to\infty$ is $N(0,\Gamma_0^{-1})$, and Γ_0 is the limit $(n\to\infty)$ of the Fisher information matrix.

To prove this theorem we need two lemmas.

Lemma 7. *Under the conditions of Theorem* 2,

$$\mathrm{P}\lim_{n\to\infty} U_n^{(ij)}(\theta_n^*) = \lim_{n\to\infty} \mathrm{E}(U_n^{(ij)}(\theta_0)) = \gamma_{ij}(\theta_0)$$

where $\theta_n^*=\omega\hat\theta_n+(1-\omega)\theta_0\in N_\delta(\theta_0)$, $0\le\omega\le 1$.

Proof is similar to that of Lemma 9 in [5], and so is omitted.

Lemma 8. *Under the conditions of Theorem* 2 *the limiting distribution of the vector* $(-\sqrt{n}\, U_n^{(1)}(\theta_0), \ldots, -\sqrt{n}\, U_n^{(p)}(\theta_0))'$ *is* $N(0, \Gamma_0)$.

Proof. To prove this lemma it is sufficient to show that for any non-zero vector $v=(v_1, \ldots, v_p)'$ the random variable

$$\Delta_n(\theta_0) \stackrel{\text{def}}{=} -\sqrt{n} \sum_{k=1}^{p} v_k U_n^{(k)}(\theta_0) = \tag{35}$$

$$= \frac{\sqrt{n}}{4\pi} \sum_{k=1}^{p} v_k \int_{-\pi}^{\pi} \left[\frac{\tilde{I}_n(t)}{h(t;\, \theta_0)} - 1\right] \left(\frac{\partial}{\partial \theta_k} \ln h(t;\, \theta)\right)_{\theta=\theta_0} dt =$$

$$= \frac{1}{4\pi\sqrt{n}} \iiint_{-\pi}^{\pi} G_n^{|Q_m|^2}(\lambda, t) G_n^{|Q_m|^2}(t, \mu) |Q_m(t)|^2 \frac{a(t;\, \theta_0)}{h(t;\, \theta_0)} dt\, Z^f(d\lambda)\overline{Z^f(d\mu)} -$$

$$- \frac{\sqrt{n}}{4\pi} \int_{-\pi}^{\pi} a(t;\, \theta_0)\, dt,$$

where

$$a(t;\, \theta_0) = \sum_{k=1}^{p} v_k (\partial \ln h(t;\, \theta)/\partial\theta_k)_{\theta=\theta_0},$$

has the limiting distribution $N(0, \gamma^2/2)$, where

$$\gamma^2 = (1/2\pi) \int_{-\pi}^{\pi} a^2(t;\, \theta_0)\, dt = 2v'\Gamma_0 v.$$

Let us denote

$$\Psi_n(\lambda, \mu;\, \theta_0) = \int_{-\pi}^{\pi} G_n^{|Q_m|^2}(\lambda, t) G_n^{|Q_m|^2}(t, \mu) \frac{|Q_m(t)|^2 a(t;\, \theta_0)}{h(t;\, \theta_0)} dt. \tag{36}$$

Since the function $\Psi_n(\lambda, \mu; \theta_0)$ is Hermitian-symmetric in (λ, μ) and belongs to L_h^2, by Schmidt's theorem (see [10]) we get

$$\Psi_n(\lambda, \mu;\, \theta_0) = \sum_{j=1}^{n} \nu_j(\theta_0) \varphi_j(\lambda;\, \theta_0) \overline{\varphi_j(\mu;\, \theta_0)}, \tag{37}$$

where $\nu_j(\theta_0)$, $j=\overline{1, n}$, is the sequence of the eigen-values and $\varphi_j(\lambda; \theta_0)$, $j=\overline{1, n}$, is the sequence of the orthonormal eigen-functions of the operator $\Phi_n(a_0/h_0)\Phi_n h_0$. The latter is an integral operator in L_h^2 generated by the kernel $\Psi_n(\lambda, \mu; \theta_0)$. Now from (36) and (37) we have

$$\eta_n(\theta_0) \stackrel{\text{def}}{=} (1/4\pi\sqrt{n}) \iint_{-\pi}^{\pi} \Psi_n(\lambda, \mu;\, \theta_0) Z^f(d\lambda)\overline{Z^f(d\mu)} = \tag{38}$$

$$= (1/4\pi\sqrt{n}) \sum_{j=1}^{n} \nu_j(\theta_0) \iint_{-\pi}^{\pi} \varphi_j(\lambda;\, \theta_0)\overline{\varphi_j(\mu;\, \theta_0)} Z^f(d\lambda)\overline{Z^f(d\mu)} =$$

$$= (1/4\pi\sqrt{n}) \sum_{j=1}^{n} \nu_j(\theta_0) \left| \int_{-\pi}^{\pi} \varphi_j(\lambda;\, \theta_0) Z^f(d\lambda)\right|^2 = (1/4\pi\sqrt{n}) \sum_{j=1}^{n} \nu_j(\theta_0) y_j^2(\theta_0),$$

where $y_j(\theta_0)=\int_{-\pi}^{\pi}\varphi_j(\lambda;\theta_0)Z^f(d\lambda),\ j=\overline{1,n}$, is a sequence of independent identically $N(0,1)$ distributed random variables.

It is well known that the characteristic function $\varphi_{\eta_n}(\alpha)$ of the random variable $\eta_n(\theta_0)$ has the form (see [7], [8])

$$\varphi_{\eta_n}(\alpha)=\prod_{j=1}^{n}\left(1-(i\alpha v_j(\theta_0)/4\pi\sqrt{n})\right)^{-1/2}. \tag{39}$$

Therefore from (35), (38) and (39) it follows that the characteristic function $\varphi_{\Delta_n}(\alpha)$ of the random variable $\Delta_n(\theta_0)$ has the form

$$\varphi_{\Delta_n}(\alpha)=\exp\left\{-(i\alpha\sqrt{n}/4\pi)\int_{-\pi}^{\pi}a(t;\theta_0)\,dt\right\}\prod_{j=1}^{n}\left(1-(i\alpha v_j(\theta_0)/4\pi\sqrt{n})\right)^{-1/2}$$

and hence

(40)

$$\ln\varphi_{\Delta_n}(\alpha)=-(1/2)\sum_{j=1}^{n}\ln\left(1-(i\alpha v_j(\theta_0)/4\pi\sqrt{n})\right)-(i\alpha\sqrt{n}/4\pi)\int_{-\pi}^{\pi}a(t;\theta_0)\,dt.$$

Using the inequalities (16), (17) and Lemma 1 in [5] it is easy to show that

$$(1/\sqrt{n})|\Phi_n(a_0/h_0)\Phi_nh_0|_h\to0,\quad n\to\infty$$
$$\sup_n(1/\sqrt{n})\|\Phi_n(a_0/h_0)\Phi_nh_0\|_h<\infty.$$

Therefore by Lemma 3 we have

$$\lim_{n\to\infty}\left[\sum_{j=1}^{n}\left\{\ln\left(1-\frac{i\alpha v_j(\theta_0)}{4\pi\sqrt{n}}\right)+\frac{i\alpha v_j(\theta_0)}{4\pi\sqrt{n}}-\frac{1}{2}\left(\frac{i\alpha v_j(\theta_0)}{4\pi\sqrt{n}}\right)^2\right\}\right]= \tag{41}$$

$$=\lim_{n\to\infty}\left[\ln\det\left(E_n-\Phi_n\frac{i\alpha a_0}{4\pi\sqrt{n}\,h_0}\Phi_nh_0\right)+\right.$$

$$\left.+\operatorname{tr}\left(\Phi_n\frac{i\alpha a_0}{4\pi h_0\sqrt{n}}\Phi_nh_0\right)-\frac{1}{2}\left\|\Phi_n\frac{i\alpha a_0}{4\pi h_0\sqrt{n}}\Phi_nh_0\right\|_h^2\right]=0.$$

Further, by Lemma 1 we have

$$\lim_{n\to\infty}\left[\operatorname{tr}\left(\Phi_n\frac{i\alpha a_0}{4\pi h_0\sqrt{n}}\Phi_nh_0\right)-\frac{i\alpha\sqrt{n}}{4\pi}\int_{-\pi}^{\pi}a(t;\theta_0)\,dt\right]=0,$$

and by Lemma 2,

$$\lim_{n\to\infty}\left[\frac{1}{2}\left\|\Phi_n\frac{i\alpha a_0}{4\pi h_0\sqrt{n}}\Phi_nh_0\right\|_h^2-\frac{\alpha^2}{4\pi}\int_{-\pi}^{\pi}a^2(t;\theta_0)\,dt\right]=0.$$

Therefore from (40) and (41) we obtain

$$\lim_{n\to\infty} \ln \varphi_{\Delta_n}(\alpha) = -(\alpha^2/4\pi) \int_{-\pi}^{\pi} a^2(t;\ \theta_0)\, dt.$$

Thus the random variable $\Delta_n(\theta_0)$ has the limiting distribution $N(0, \gamma^2/2)$.

Proof of Theorem 2. Since $U_n^{(j)}(\hat\theta_n)=0$, $j=\overline{1,p}$, by the mean value theorem we have

$$0 = U_n^{(j)}(\hat\theta_n) = U_n^{(j)}(\theta_0) + \sum_{i=1}^{p} (\hat\theta_{ni} - \theta_{0i})\, U_n^{(ij)}(\theta_n^*), \tag{42}$$

where $\theta_n^* = \omega\hat\theta_n + (1-\omega)\theta_0 \in N_\delta(\theta_0)$, $0 \leqq \omega \leqq 1$. The relation (42) may be rewritten as

$$-\sqrt{n}\, U_n^{(j)}(\theta_0) = \sum_{t=1}^{p} \sqrt{n}\, (\hat\theta_{ni} - \theta_{0i}) U_n^{(ij)}(\theta_n^*). \tag{43}$$

Now by Lemma 7

$$\mathrm{P}\lim_{n\to\infty} U_n^{(ij)}(\theta_n^*) = \gamma_{ij}(\theta_0),$$

and by Lemma 8 the random vector $(-\sqrt{n}\ U_n^{(1)}(\theta_0), \ldots, -\sqrt{n}\ U_n^{(p)}(\theta_0))'$ has the limiting distribution $N(0, \Gamma_0)$. Therefore (43) implies that the vector $\sqrt{n}\ (\hat\theta_n - \theta_0)$ has the limiting distribution $N(0, \Gamma_0^{-1})$.

Finally, let us show that the matrix Γ_0 is the limit $(n\to\infty)$ of the Fisher information matrix. This statement follows from the following relation

$$\lim_{n\to\infty} (1/n)\mathrm{D}\Big(\sum_{i=1}^{p} v_i \partial L_n(\theta)/\partial\theta_i\Big) = \lim_{n\to\infty} (1/n)\mathrm{D}\Big(\sum_{i=1}^{p} v_i U_n^{(i)}(\theta)\Big) =$$

$$= \lim_{n\to\infty} (1/n)\,\mathrm{D}\Big((1/4\pi) \iiint_{-\pi}^{\pi} G_n^{|Q_m|^2}(\lambda, t) G_n^{|Q_m|^2}(t, \mu) |Q_m(t)|^2 \times$$

$$\times (a(t;\ \theta)/h(t;\ \theta))\, dt\, Z^f(d\lambda)\overline{Z^f(d\mu)} - (1/4\pi) \int_{-\pi}^{\pi} a(t;\ \theta)\, dt\Big) =$$

$$= \lim_{n\to\infty} (1/4\pi n) \iint_{-\pi}^{\pi} \Big| \int_{-\pi}^{\pi} G_n^{|Q_m|^2}(\lambda, t) G_n^{|Q_m|^2}(t, \mu) |Q_m(t)|^2 \times$$

$$\times (a(t;\ \theta)/h(t;\ \theta))\, dt\Big|^2 f(\lambda;\ \theta) f(\mu;\ \theta)\, d\lambda\, d\mu = (1/4\pi) \int_{-\pi}^{\pi} a^2(t;\ \theta)\, dt,$$

where, as before, $v=(v_1, \ldots, v_p)'$ is a non-zero vector and

$$a(t;\ \theta) = \sum_{i=1}^{p} v_i \partial \ln h(t;\ \theta)/\partial\theta_i.$$

4. Confidence regions for the parameter θ

The further arguments are based on the following theorem.

Theorem 3. *Let* $\hat{\theta}_n$ *be an arbitrary consistent, asymptotically normal and asymptotically efficient estimate of the parameter* θ *and let the random matrix* $\Gamma_* = \|\gamma^*_{kj}\|_{kj=\overline{1,p}}$ *be an arbitrary consistent estimate of the limit* Γ_0 *of the Fisher information matrix. Then the limiting* $(n\to\infty)$ *distribution of the statistic*

$$S_n^2 = n \sum_{ij=1}^{p} (\hat{\theta}_{ni}-\theta_{0i})(\hat{\theta}_{nj}-\theta_{0j})\gamma^*_{ij} \tag{44}$$

is the χ^2*-distribution with* p *degrees of freedom.*

Proof. It is easy to see that

$$\mathrm{P}\lim_{n\to\infty} n \sum_{ij=1}^{p} (\hat{\theta}_{ni}-\theta_{0i})(\hat{\theta}_{nj}-\theta_{0j})[\gamma^*_{ij}-\gamma_{ij}(\theta_0)] = 0.$$

Hence the limiting distribution of the statistic S_n^2 is the same as that of

$$n \sum_{ij=1}^{p} (\hat{\theta}_{ni}-\theta_{0i})(\hat{\theta}_{nj}-\theta_{0j})\gamma_{ij}(\theta_0). \tag{45}$$

Transforming the vector $\sqrt{n}\,(\hat{\theta}_n-\theta_0)$ to a vector ξ via the unitary transformation V such that the matrix $V'\Gamma_0V$ is diagonal, we obtain

$$n(\hat{\theta}_n-\theta_0)'\Gamma_0(\hat{\theta}_n-\theta_0) = \sum_{i=1}^{p} \xi_i^2/\sigma_i^2,$$

where ξ_i is the i-th component of ξ and σ_i^2 is its variance. The random variables ξ_i, $i=\overline{1,p}$, converge in probability to independent normal random variables with mean 0 and variance σ_i^2, $i=\overline{1,p}$. Therefore the random variable (45) and hence the statistic S_n^2 has the limiting χ^2-distribution with p degrees of freedom.

Thus we have shown that for every interval $[\alpha, \beta]$ the relation

$$\alpha < n(\hat{\theta}_n-\theta_0)'\Gamma_*(\theta_n-\theta_0) < \beta \tag{46}$$

has limiting probability $\int_\alpha^\beta \chi_p^2(x)\,dx$. If α and β are chosen so that

$$\int_\alpha^\beta \chi_p^2(x)\,dx = 1-\varepsilon, \quad \varepsilon > 0,$$

then the set of values of θ satisfying (46) will be a confidence region for θ_0 with asymptotic confidence level ε.

References

[1] N. I. Ahiezer—I. M. Glazman, *The theory of linear operators in Hilbert space*, Nauka (Moscow, 1966). (Russian)

[2] R. B. Davies, Asymptotic inference in stationary Gaussian timeseries, *Adv. Appl. Probab.*, 5 (1973), 469—497.

[3] K. O. Dzhaparidze, *Estimation of parameters and testing hypotheses in spectral analysis of stationary timeseries*, University Press (Tbilisi, 1981). (Russian)

[4] K. O. Dzhaparidze, An estimation of parameters of spectral density with fixed zeros, *Teoria Verojat. i Primen.*, 22 (1977), 729—748. (Russian)

[5] M. S. Ginovian, On the asymptotical estimation of the maximum likelihood of parameters of the spectrum of a stationary Gaussian timeseries, in: *Limit theorems in probability and statistics*, Vol. I. (P. Révész, ed.) Coll. Math. Soc. J. Bolyai, 36, North-Holland (Amsterdam, 1984), 457—497.

[6] M. S. Ginovian, The asymptotic behaviour of the likelihood function involving polynomial zeros of spectral density, *Zapiski Nauchn. Semin. LOMI*, 108 (1981), 5—21. (Russian)

[7] U. Grenander—G. Szegő, *Toeplitz forms and their applications*, University of California Press (Berkeley, 1958).

[8] I. A. Ibrahimov, On estimation of the spectral function of a stationary Gaussian process, *Teoria Verojat. i Primen.*, 8 (1963), 391—430. (Russian)

[9] I. A. Ibrahimov—Yu. A. Rozanov, *Gaussian stochastic processes*, Nauka (Moscow, 1970). (Russian)

[10] F. Riesz—B. Sz.-Nagy, *Leçons d'analyse fonctionnelle*, Akadémiai Kiadó (Budapest, 1972).

[11] A. M. Walker, Asymptotic properties of least-squares estimation of parameters of the spectrum of a stationary non-deterministic time-series, *J. Austral. Math. Soc.*, 4 (1964), 363—384.

[12] P. Whittle, Estimation and information in stationary time-series, *Arch. Math.*, 2 (1953), 423—434.

INSTITUTE OF MATHEMATICS OF THE
ACADEMY OF SCIENCES OF THE ARMENIAN SSR
MARSHAL BAGRAMIAN ST. 24 B
375019 — YEREVAN, ARM. SSR, USSR

Acta Sci. Math, 50 (1986), 183—189

Decompositions of completely bounded maps

TADASI HURUYA

1. Introduction. Let $\varphi\colon A\to B$ be a bounded linear map between C^*-algebras, and let $\varphi\otimes \mathrm{id}_n\colon A\otimes M_n\to B\otimes M_n$ be the associated maps ($n=1, 2, \dots$), where M_n is the full matrix algebra of order n. The map φ is said to be completely positive if each $\varphi\otimes\mathrm{id}_n$ is positive, and completely bounded if $\|\varphi\|_{\mathrm{cb}}\equiv\sup_n\|\varphi\otimes\mathrm{id}_n\|<\infty$; in this case $\|\varphi\|_{\mathrm{cb}}$ is called the completely bounded norm of φ. It is known that a completely positive map φ is completely bounded and $\|\varphi\|_{\mathrm{cb}}=\|\varphi\|$. A linear map $\varphi\colon A\to B$ between C^*-algebras is said to have a positive (resp. completely positive) decomposition if φ can be written as a linear combination of positive (resp. completely positive) linear maps.

A C^*-algebra A is injective if and only if for any C^*-algebra B such that $B\supseteqq A$, there exists a projection of B onto A of norm one [2; Theorem 5.3]. WITTSTOCK [15] proved that every completely bounded map of a unital C^*-algebra into an injective C^*-algebra has a completely positive decomposition (see, also, [8]). In [3] we proved, as a limited converse of Wittstock's theorem, that given a separable C^*-algebra B, every bounded linear map of any C^*-algebra into B has a positive decomposition if and only if B is finite-dimensional, namely, injective. In this paper, we show that given a separable unital C^*-algebra B, every completely bounded map of any unital C^*-algebra into B has a completely positive decomposition if and only if B is finite-dimensional, namely, injective. We also prove that if A and B are separable, infinite-dimensional, unital C^*-algebras and A contains a self-adjoint element such that the set of limit points of its spectrum is infinite, then the span of positive linear maps of A into B is nowhere dense in the Banach space of all bounded linear maps of A into B.

Throughout this paper, it is assumed that all C^*-algebras are unital. If S is a compact Hausdorff space, we denote by $C(S)$ the C^*-algebra of continuous functions on S. We mean by αN the one-point compactification of the set N of positive integers with the point ∞ at infinity.

Received May 20, 1983.

We remark that some of the results of this paper were announced in [12].

The author would like to thank Professor J. Tomiyama for his interest and encouragement.

2. Bounded linear maps between commutative C^*-algebras. For each n in N let $X_n=\{x_{n,1},\cdots,x_{n,m},\cdots,x_{n,\infty}\}$ and $Y_n=\{y_{n,1},\cdots,y_{n,m},\ldots\}$. Denote by X the one-point compactification of the topological sum of the sequence $\{X_n\}_{n=1}^{\infty}$ of copies αN; denote by Y the one-point compactification of the sequence $\{Y_n\}_{n=1}^{\infty}$ of copies N with the point y_∞ at infinity. The space Y is homeomorphic to αN.

From now on, we use X, Y, X_n, $x_{n,m}$, $x_{n,\infty}$, $y_{n,m}$ and y_∞ in the above situation. We consider linear maps of $C(X)$ into $C(Y)$. Proofs of Lemmas 2 and 3 are based on an idea due to Kaplan [4] and Tsui [13].

We first recall Tsui's example [13; 1.3.4, Example II].

Lemma 1. *Let Φ: $C(X)\to C(Y)$ be the self-adjoint linear map defined by $\Phi(f)(y_{n,m})=f(x_{n,m})-f(x_{n,m+1})$ and $\Phi(f)(y_\infty)=0$. Then Φ has no positive decomposition.*

Lemma 2. *For any positive integer k, there exists a self-adjoint linear map Ψ_k: $C(X)\to C(Y)$ with $\|\Psi_k\|=1$ satisfying the following properties.*

(1) Ψ_k has a positive decomposition.

(2) If Ψ_k is decomposed as the difference of two positive linear maps Ψ^+, Ψ^-, then $\|\Psi^+\|\geqq k/2$ and $\|\Psi^-\|\geqq k/2$.

Proof. We define the map Ψ_k: $C(X)\to C(Y)$ by

$$\Psi_k(f)(y_{n,m})=(1/2)(f(x_{n,m})-f(x_{n,m+1}))\quad\text{if}\quad n\leqq k,$$
$$\Psi_k(f)(y_{n,m})=0\quad\text{if}\quad n>k,$$
$$\Psi_k(f)(y_\infty)=0.$$

It is easy to check that $\Psi_k(f)$ is continuous on Y and $\|\Psi_k\|=1$.

(1) We define k positive linear maps ψ_i: $C(X)\to C(Y)$ by

$$\psi_i(f)(y_{n,m})=0\quad\text{if}\quad n+m\leqq i,$$
$$\psi_i(f)(y_{n,m})=(1/2)f(x_{i,n+m-i})\quad\text{if}\quad n+m>i,$$
$$\psi_i(f)(y_\infty)=(1/2)f(x_{i,\infty}).$$

Since $\psi_i(f)(y_{i,m})=(1/2)f(x_{i,m})$, we have $\psi_1+\ldots+\psi_k\geqq\Psi_k$, and hence Ψ_k has a positive decomposition.

(2) For a linear map ψ: $C(X)\to C(Y)$, let $\psi_{(n,m)}$ denote the linear functional on $C(X)$ defined by $\psi_{(n,m)}(f)=\psi(f)(y_{n,m})$. If $n\leqq k$, then $(1/2)\delta(x_{n,m})\leqq\Psi^+_{(n,m)}$ and

$(1/2)\delta(x_{n,m+1}) \leqq \Psi^-_{(n,m)}$, where $\delta(x)$ denotes the point measure at x. Let e_n be the characteristic function of the subset X_n of X. Then

$$1/2 = (1/2)\delta(x_{n,m})(e_n) \leqq \Psi^+_{(n,m)}(e_n) = \Psi^+(e_n)(y_{n,m}),$$

$$1/2 = (1/2)\delta(x_{n,m+1})(e_n) \leqq \Psi^-_{(n,m)}(e_n) = \Psi^-(e_n)(y_{n,m}).$$

If $j=+, -$, we have

$$1/2 \leqq \lim_m \Psi^j(e_n)(y_{n,m}) = \Psi^j(e_n)(y_\infty).$$

Therefore,

$$\Psi^j(1)(y_\infty) \geqq \sum_{n=1}^{k} \Psi^j(e_n)(y_\infty) \geqq k/2,$$

so that $\|\Psi^+\| \geqq k/2$ and $\|\Psi^-\| \geqq k/2$.

Lemma 3. *With Φ as in Lemma* 1, *if φ is a bounded linear map of $C(X)$ into $C(Y)$ satisfying $\|\varphi - \Phi\| < 1/2$, then φ has no positive decomposition.*

Proof. Suppose that φ has a positive decomposition. Then the self-adjoint part τ of φ has a positive decomposition $\tau = \tau^+ - \tau^-$. For a linear map ψ: $C(X) \to C(Y)$, as in Lemma 2, we define the linear functional $\psi_{(n,m)}$ on $C(X)$ by $\psi_{(n,m)}(f) = \psi(f)(y_{n,m})$. Since X is countable, $\tau^j_{(n,m)}$ can be written as

$$\tau^j_{(n,m)} = \sum_{x \in X} \beta^j(x)\delta(x), \quad 0 \leqq \beta^j(x) \in \mathbf{R}, \quad j = +, -,$$

where $\delta(x)$ denotes the point measure at x. We then have

$$1/2 > \|\varphi - \Phi\| \geqq \|\tau - \Phi\| \geqq \|\tau_{(n,m)} - \Phi_{(n,m)}\| =$$
$$= \Big\|\sum_{x \in X} \beta^+(x)\delta(x) - \sum_{x \in X} \beta^-(x)\delta(x) - \delta(x_{n,m}) + \delta(x_{n,m+1})\Big\| \geqq$$
$$\geqq |\beta^+(x_{n,m}) - \beta^-(x_{n,m}) - 1|.$$

Hence $(1/2)\delta(x_{n,m}) \leqq \tau^+_{(n,m)}$. Let e_n be the characteristic function of the subset X_n of X. Then

$$1/2 = (1/2)\delta(x_{n,m})(e_n) \leqq \tau^+_{(n,m)}(e_n) = \tau^+(e_n)(y_{n,m}),$$

so that $1/2 \leqq \lim_m \tau^+(e_n)(y_{n,m}) = \tau^+(e_n)(y_\infty)$. Therefore

$$\tau^+(1)(y_\infty) \geqq \tau^+\Big(\sum_{n=1}^{k} e_n\Big)(y_\infty) \geqq k/2$$

for any positive integer k. This implies the unboundedness of τ^+.

3. The main results. We recall that X and Y are the one-point compactifications of the topological sums of sequences of copies αN and N, respectively. In order to extend maps obtained in Section 2 to non-commutative C^*-algebras, we construct completely positive maps with range algebras $C(X)$, $C(Y)$.

Lemma 4. *If a separable C^*-algebra A contains a self-adjoint element a such that the set of limit points of the spectrum of a is infinite, then there exist unital completely positive maps π_A: $A \to C(X)$ and ν_A: $C(X) \to A$ such that $\pi_A \circ \nu_A$ is the identity map on $C(X)$.*

Proof. Let S denote the spectrum of a. Since S is a compact subset of real numbers, choose a point s_∞ and a sequence $\{s_n\}_{n=1}^\infty$ of limit points of S such that $3|s_\infty - s_{n+1}| < |s_\infty - s_n|$ for all n. For each n take a sequence $\{s_{n,i}\}_{i=1}^\infty$ of distinct points of S such that $3|s_n - s_{n,i}| < |s_\infty - s_n|$ for all i and $\lim_i s_{n,i} = s_n$. Put $S_n = \{s_{n,1}, \ldots, s_{n,m}, \ldots, s_n\}$ and $\tilde{S} = \{s_\infty\} \cup (\bigcup_{n=1}^\infty S_n)$. If $s \in \tilde{S}$, we choose a state g_s on A such that $g_s(f) = f(s)$ for all f in $C(S)$ because $C(S)$ is the C^*-subalgebra generated by a and 1. We then define the positive linear map π of A into the C^*-algebra of all bounded functions on $\tilde{S}$ by $\pi(b)(s) = g_s(b)$ for s in $\tilde{S}$ and b in A. Since A is separable, so is the C^*-subalgebra $C^*(\pi(A))$ generated by $\pi(A)$. There exists a compact metric space T with metric d such that $C(T) = C^*(\pi(A))$. Then $\tilde{S}$ is canonically regarded as a subset of T. For each n let t_n be a limit point of the subset S_n of T and choose a subsequence $\{\tilde{s}_{n,i}\}_{i=1}^\infty$ of $\{s_{n,i}\}_{i=1}^\infty$ such that $\lim_i \tilde{s}_{n,i} = t_n$. If $f \in C(S)$, then

$$\pi(f)(t_n) = \lim_i \pi(f)(\tilde{s}_{n,i}) = \lim_i f(\tilde{s}_{n,i}) = f(s_n).$$

Hence $t_n \neq t_m$ if $n \neq m$.

We again choose a point t_∞ and a subsequence $\{t_{h(n)}\}_{n=1}^\infty$ of $\{t_n\}_{n=1}^\infty$ such that $3d(t_\infty, t_{h(n+1)}) < d(t_\infty, t_{h(n)})$. For each n take a subsequence $\{t_{h(n),i}\}_{i=1}^\infty$ of $\{\tilde{s}_{h(n),i}\}_{i=1}^\infty$ such that $3d(t_{h(n)}, t_{h(n),i}) < d(t_\infty, t_{h(n)})$. Put $T_n = \{t_{h(n),1}, \ldots, t_{h(n),m}, \ldots, t_{h(n)}\}$ and $\tilde{X} = \{t_\infty\} \cup (\bigcup_{n=1}^\infty T_n)$. By its construction, $\tilde{X}$ is canonically regarded as the space X.

Let φ: $X = \tilde{X} \to \tilde{S} \subseteqq S$ be defined by

$$\varphi(t_{h(n),i}) = t_{h(n),i}, \quad \varphi(t_{h(n)}) = s_{h(n)}, \quad \varphi(t_\infty) = s_\infty.$$

For f in $C(S)$, $\pi(f)(t_{h(n)}) = f(s_{h(n)})$ and $\pi(f)(t_\infty) = f(s_\infty)$. The map φ is one-to-one and continuous. Then there exists, by [1; Theorem 3.11], a unital positive linear map ν_A: $C(X) \to C(S) \subseteqq A$ such that $\nu_A(f) \circ \varphi = f$ for all f in $C(X)$.

Define the unital positive linear map π_A: $A \to C(X)$ by $\pi_A(b) = \pi(b)|_X$, the restriction to $X = \tilde{X}$ of $\pi(b)$. Then $\pi_A(f) = f \circ \varphi$ for all f in $C(S)$. Hence $\pi_A \circ \nu_A$

is the identity map on $C(X)$. Both π_A and ν_A are completely positive [10; Chapter IV, Corollary 3.5].

Lemma 5. *If B is a separable, infinite-dimensional C^*-algebra, then there exist unital completely positive maps $\pi_B\colon B\to C(Y)$ and $\nu_B\colon C(Y)\to B$ such that $\pi_B\circ\nu_B$ is the identity map on $C(Y)$.*

Proof. There exists a self-adjoint element a in B with infinite spectrum S [7]. Denote by $C^*(a, 1)$ the C^*-subalgebra generated by a and 1. Then $C(S)=C^*(a, 1)$. Since S is a compact metrizable space, we choose a point s_∞ and a sequence $\{s_n\}_{n=1}^\infty$ of distinct points in S with $\lim_n s_n=s_\infty$ and $\{a_n\}_{n=1}^\infty$ of $C(S)$ such that $a_n(s_n)=1$, $0\leqq a_n\leqq 1$ and $a_p a_q=0$ for $p\neq q$.

Put $\tilde{S}=\{s_1, \dots, s_n, \dots, s_\infty\}$. If $s\in\tilde{S}$, we take a state g_s on B such that $g_s(f)=f(s)$ for all f in $C(S)$. We define the unital positive linear map π of B into the C^*-algebra of all bounded functions on $\tilde{S}$ by $\pi(b)(s)=g_s(b)$. Since B is separable, so is the C^*-subalgebra $C^*(\pi(B))$ generated by $\pi(B)$. There then exists a compact metrizable space T such that $C(T)=C^*(\pi(B))$. Then $\tilde{S}$ is canonically regarded as a subset of T.

We choose a point $s_{h(\infty)}$ in T and a subsequence $\{s_{h(n)}\}_{n=1}^\infty$ of $\{s_n\}_{n=1}^\infty$ such that $\lim_n s_{h(n)}=s_{h(\infty)}$. Put $\tilde{Y}=\{s_{h(1)}, \dots, s_{h(n)}, \dots, s_{h(\infty)}\}$. Then $\tilde{Y}$ is canonically regarded as the space Y because $\tilde{Y}$ is homeomorphic to αN.

We define the unital positive linear map $\pi_B\colon B\to C(Y)$ by $\pi_B(b)=\pi(b)|_Y$, the restriction to $Y=\tilde{Y}$ of $\pi(b)$. We also define the unital positive linear map $\nu_B\colon C(Y)\to C(S)\subseteqq B$ by

$$\nu_B(b) = \sum_{n=1}^{\infty} [b(s_{h(n)})-b(s_{h(\infty)})]a_{h(n)}+b(s_{h(\infty)})1.$$

Then $\pi_B\circ\nu_B$ is the identity map on $C(Y)$ and both π_B and ν_B are completely positive [10; Chapter IV, Corollary 3.5].

Theorem 6. *Let A and B be separable, infinite-dimensional C^*-algebras. Assume that A contains a self-adjoint element a such that the set of limit points of the spectrum of a is infinite. Then*

(1) *There exists a completely bounded map $\tilde{\Phi}\colon A\to B$ such that each bounded linear map $\varphi\colon A\to B$ with $\|\varphi-\tilde{\Phi}\|<1/2$ has no positive decomposition.*

(2) *There exists a self-adjoint linear map $\tilde{\Psi}\colon A\to B$ having a completely positive decomposition such that for any positive decomposition $\tilde{\Psi}=\tilde{\Psi}^+-\tilde{\Psi}^-$ we have $\|\tilde{\Psi}^+\|>\|\tilde{\Psi}\|_{cb}$ and $\|\tilde{\Psi}^-\|>\|\tilde{\Psi}\|_{cb}$.*

Proof. We use maps Φ, Ψ_4, π_A, ν_A, π_B and ν_B constructed in Lemmas 1, 2, 4 and 5.

(1) We put $\tilde{\Phi}=\nu_B\circ\Phi\circ\pi_A$. By [6; Lemma 1], $\tilde{\Phi}$ is completely bounded. Then

$$\|\pi_B\circ\varphi\circ\nu_A-\Phi\| = \|\pi_B\circ\varphi\circ\nu_A-\pi_B\circ\tilde{\Phi}\circ\nu_A\| \leqq \|\varphi-\tilde{\Phi}\| < 1/2.$$

By Lemma 3, $\pi_B\circ\varphi\circ\nu_A$ has no positive decomposition. If φ has a positive decomposition, so does $\pi_B\circ\varphi\circ\nu_A$. This is a contradiction.

(2) We put $\tilde{\Psi}=\nu_B\circ\Psi_4\circ\pi_A$. By Lemma 2 and [10; Chapter IV, Corollary 3.5], $\tilde{\Psi}$ is a self-adjoint linear map of A into B having a completely positive decomposition and

$$1 = \|\Psi_4\| = \|\pi_B\circ\tilde{\Psi}\circ\nu_A\| \leqq \|\tilde{\Psi}\|_{cb} = \|\nu_B\circ\Psi_4\circ\pi_A\|_{cb} \leqq \|\Psi_4\|_{cb}=\|\Psi_4\|,$$

where the last equality follows from [6; Lemma 1]. Hence $\|\tilde{\Psi}\|_{cb}=1$. If $\tilde{\Psi}$ has a positive decomposition $\tilde{\Psi}=\tilde{\Psi}^+-\tilde{\Psi}^-$, then Ψ_4 has a positive decomposition $\Psi_4= =\pi_B\circ\tilde{\Psi}^+\circ\nu_A-\pi_B\circ\tilde{\Psi}^-\circ\nu_A$. By Lemma 2 we have

$$\|\tilde{\Psi}^+\| \geqq \|\pi_B\circ\tilde{\Psi}^+\circ\nu_A\| \geqq 4/2 > \|\tilde{\Psi}\|_{cb},$$

and similarly,

$$\|\tilde{\Psi}^-\| \geqq 2 > \|\tilde{\Psi}\|_{cb}.$$

Remark 7. Let A_1 and B_1 be C^*-algebras. Suppose that there exist unital completely positive maps $\pi_1: A_1\to C(X)$, $\nu_1: C(X)\to A_1$, $\pi_2: B_1\to C(Y)$, $\nu_2: C(Y)\to B_1$ such that $\pi_1\circ\nu_1$ and $\pi_2\circ\nu_2$ are the identity maps on $C(X)$ and $C(Y)$, respectively. If we replace A and B by A_1 and B_1, Theorem 6 remains true from the same argument in its proof (cf. [9; Theorem 2.6]).

We recall that the set of self-adjoint elements of an injective C^*-algebra is conditionally complete [11; Theorem 7.1]. Hence a separable C^*-algebra A is injective if and only if A is finite-dimensional.

Corollary 8. *Let B be a separable C^*-algebra. The following statements are equivalent;*

(1) *B is injective;*

(2) *Every completely bounded map of any C^*-algebra into B has a completely positive decomposition;*

(3) *Every linear map φ having a completely positive decomposition of any C^*-algebra into B has a completely positive decomposition such that $\varphi=\varphi_1-\varphi_2+i(\varphi_3-\varphi_4)$ with $\|\varphi_i\|\leqq\|\varphi\|_{cb}$ $(i=1, \cdots, 4)$.*

Proof. By [15; Satz 4.5] we have (1)⇒(2) and (1)⇒(3). Combining the above remark about injective, separable C^*-algebras with Theorem 6, we see that (2)⇒(1) and (3)⇒(1).

In the category of partially ordered Banach spaces, Wickstead [14, Theorem 3.15] obtained a result similar to Corollary 8.

Addition. After this paper was written, the author discovered an example of a non-injective, non-separable C^*-algebra B such that every completely bounded map of any C^*-algebra into B has a completely positive decomposition [16].

References

[1] W. G. BADE, *The Banach space $C(S)$*, Aarhus University Lecture Notes Series, No. 26, 1971.

[2] E. G. EFFROS and E. C. LANCE, Tensor products of operator algebras, *Adv. in Math.*, **25** (1977), 1—34.

[3] T. HURUYA and J. TOMIYAMA, Completely bounded maps of C^*-algebras, *J. Operator Theory*, to appear.

[4] S. KAPLAN, An example in the space of bounded operators from $C(X)$ to $C(Y)$, *Proc. Amer. Math. Soc.*, **38** (1973), 595—597.

[5] H. E. LACEY, *The isometric theory of classical Banach spaces*, Springer-Verlag (Berlin—Heidelberg—New York, 1974).

[6] R. I. LOEBL, Contractive linear maps on C^*-algebras, *Michigan Math. J.*, **22** (1975), 361—366.

[7] T. OGASAWARA, Finite dimensionality of certain Banach algebras, *J. Sci. Hiroshima Univ.*, **17** (1954), 359—364.

[8] V. I. PAULSEN, Completely bounded maps on C^*-algebras and invariant operator ranges, *Proc. Amer. Math. Soc.*, **86** (1982), 91—96.

[9] R. R. SMITH, Completely bounded maps between C^*-algebras, *J. London Math. Soc.*, (2) **27** (1983), 157—166.

[10] M. TAKESAKI, *Theory of operator algebras. I*, Springer-Verlag (Berlin—Heidelberg—New York, 1979).

[11] J. TOMIYAMA, *Tensor products and projections of norm one in von Neumann algebras*, Lecture Notes (mimeographed), University of Copenhagen, 1970.

[12] J. TOMIYAMA, Recent development of the theory of completely bounded maps between C^*-algebras, to appear.

[13] S.-K. J. TSUI, Decompositions of linear maps, *Trans. Amer. Math. Soc.*, **230** (1977), 87—112.

[14] A. W. WICKSTEAD, Spaces of linear operators between partially ordered Banach spaces, *Proc. London Math. Soc.*, (3) **28** (1974), 141—158.

[15] G. WITTSTOCK, Ein operatorwertiger Hahn—Banach Satz, *J. Funct. Anal.*, **40** (1981), 127—150.

[16] T. HURUYA, Linear maps between certain non-separable C^*-algebras, to appear.

FACULTY OF EDUCATION
NIIGITA UNIVERSITY
NIIGITA 950-21, JAPAN

Acta Sci. Math., 50 (1986), 191—211

Characterization of operators of class C_0 and a formula for their minimal function

AHARON ATZMON*)

Introduction

The class of operators C_0 was introduced in 1964 by Sz.-NAGY and C. FOIAŞ [11] and consists of all completely non unitary contractions on a complex Hilbert space which are annihilated by a non-identically zero function in H^∞ of the unit disc. Among the annihilating H^∞ functions of a contraction of class C_0, there exists an inner function which divides (in H^∞) all others ([11] or [12, p. 124]). This inner function is determined up to a constant factor of modulus 1, and is called the minimal function of the contractions. For results concerning the structure of contractions of class C_0 we refer to [3], [13] and [14].

The first characterization of contractions of class C_0 was given by Sz.-Nagy and Foiaş in terms of an algebraic condition on the characteristic operator function [12, p. 265]. Using this characterization, J. DAZORD [4] obtained a characterization of C_0 operators in terms of a growth condition on their resolvent, which however is of an implicit form and is difficult to verify. (See Corollary 4.2.)

In this paper we give a characterization of C_0 operators in terms of an explicit growth condition on their resolvent, and establish a formula for the associated minimal function, also in terms of the resolvent (Theorem 1.1). A similar characterization of C_0 operators whose spectrum is a thin set in a certain sense, is given in [2]. (See Section 7.)

The above mentioned characterization and formula for the minimal function can also be expressed in terms of the characteristic operator function (Theorem 5.3). The interest in obtaining such a result was pointed out by R. G. DOUGLAS [3, p. 190].

Our exposition is self contained in the sense that the concepts from operator

Received October 28, 1983.

*) This research was supported by the Fund for the Promotion of Research at the Technion — Israel Institute of Technology.

theory that are used in the proofs of the theorems, are essentially those which appear in their statement. So for example, except for Section 5, we do not use in our proofs the characteristic operator functions or functional model of Sz.-Nagy and Foiaş. Although resulting in longer proofs, this approach seems to be of interest and also leads to new proofs of the characterizations of operators of class C_0 given in [12, p. 265] and [4]. (See Corollary 4.2 and Corollary 5.4.) We also obtain a new proof of the existence of a minimal function for C_0 operators (Theorem 1.1).

The contents of the paper are as follows: In Section 1 we introduce the concept of meromorphic vector function of bounded α-characteristic and state our main result. In Section 2 we prove some preliminary results which are needed for the proof of the main result. Section 3, which is the principal part of the paper, is devoted to the study of contractions with resolvent of bounded 1-characteristic. To every such contraction T we associate a function φ_T in H^∞, which is expressed in terms of the resolvent of T, and is a minimal function of T in the case that T is of class C_0. We also characterize in this section the resolvents of operators in this class, and prove the invariance of the class under certain Möbius transformations. In Section 4 we present the proof of our main result and obtain as a Corollary the result of Dazord [4]. In Section 5 we characterize contractions T whose characteristic function θ_T has a scalar multiple, and express our main result in terms of θ_T. In Section 6 we characterize contractions of class D_0, that is, contractions which are annihilated by a non-identically zero function in the disc algebra, and give the general form of an annihilating function of such a contraction. Finally in Section 7, we consider contractions with resolvent of bounded α-characteristic for some $0 \leqq \alpha < 1$, and prove that they are of class D_0, and have (in a certain sense) a thin spectrum.

The basic notions and facts concerning the Banach algebra H^∞ and the functional calculus of Sz.-Nagy and Foiaş for completely non unitary contractions, will be used freely in the sequel without giving always an explicit reference. For H^∞ we refer to [7] or [9] and for the functional calculus to [12, Chapter III].

1. Definitions and main result

Throughout this paper, $\mathscr{H}$ will denote a complex Hilbert space and $\mathscr{L}(\mathscr{H})$ the algebra of all bounded linear operators on $\mathscr{H}$. For an operator T in $\mathscr{L}(\mathscr{H})$ we shall denote by $\sigma(T)$ its spectrum, by $\varrho(T)$ its resolvent set, and by $R_T(\lambda)$ its resolvent, $(\lambda I-T)^{-1}$, $\lambda\in\varrho(T)$. We shall also denote by L_T, the operator function defined by: $L_T(\lambda)=(I-\bar{\lambda}T)R_T(\lambda)$, $\lambda\in\varrho(T)$. The term contraction will mean in the sequel an operator T in $\mathscr{L}(\mathscr{H})$ such that $\|T\|\leqq 1$.

The open unit disc $\{\lambda\in\mathbf{C}\colon |\lambda|<1\}$ will be denoted by D and the unit circle $\{\lambda\in\mathbf{C}\colon |\lambda|=1\}$ by Γ.

If X is a complex Banach space and F is an X-valued meromorphic function on D, we shall denote for every $0 \leqq t < 1$ by $n(t, F)$ the number of poles of F in the disc $\{\lambda \in \mathbf{C}: |\lambda| \leqq t\}$ (counting multiplicity), and for every $\alpha \geqq 0$ and $0 \leqq r < 1$ we set

$$N(r, F) = \int_0^r \big((n(t, F) - n(0, F))/t\big)\, dt + n(0, F) \log r$$

and

$$m_\alpha(r, F) = (1/2\pi) \int_0^{2\pi} \log^+ \|(1-r)^\alpha F(re^{i\theta})\|\, d\theta$$

(where for $a \geqq 0$, $\log^+ a = \max\{\log a, 0\}$).

We define the *α-characteristic* of an X-valued meromorphic function on D to be the function

$$T_\alpha(F, r) = m_\alpha(F, r) + N(F, r), \quad 0 \leqq r < 1.$$

If $\sup\limits_{0 \leqq r < 1} T_\alpha(F, r) < \infty$, then we say that F is of *bounded α-characteristic.*

The set of all X-valued meromorphic functions on D of bounded α-characteristic, will be denoted by $N_\alpha(X)$. The elements in $N_0(X)$ are called functions of *bounded characteristic.* For $X = \mathbf{C}$ this is the classical definition of R. Nevanlinna [15]. Vector valued functions of bounded characteristic are considered in [2].

To simplify notations we shall denote in the sequel the set $N_\alpha(\mathscr{L}(\mathscr{H}))$ by N_α; and if T is a contraction such that the operator function $\lambda \to R_T(\lambda)$, $\lambda \in \varrho(T) \cap D$, is meromorphic on D and is in N_α, we shall say briefly that R_T is in N_α.

We recall that a contraction T is said to be of class $C_{0\cdot}$, if $T^n x \to 0$ as $n \to \infty$, for every x in $\mathscr{H}$ [12, p. 72].

Our main result is the following:

Theorem 1.1. *A contraction T is of class C_0, if and only if, T is of class $C_{0\cdot}$, and R_T is in N_1. Furthermore, if the last two conditions are satisfied, and $\{\lambda_1, \lambda_2, \ldots\}$ is the sequence of poles of R_T in D repeated according to multiplicity, with k of the λ_n being equal to zero, then T has a minimal function given by*

$$m_T(z) = z^k \prod_{\lambda_j \neq 0} (\bar\lambda_j/|\lambda_j|)\big((\lambda_j - z)/(1 - \bar\lambda_j z)\big) \exp(-w(z)), \quad z \in D,$$

where

$$w(z) = \lim_{\varrho \to 1-} (1/2\pi) \int_0^{2\pi} \big((e^{it} + z)/(e^{it} - z)\big) \log \|L_T(\varrho e^{it})\|\, dt, \quad z \in D,$$

or alternatively,

$$w(z) = \int_\Gamma \big((e^{it} + z)/(e^{it} - z)\big)\, d\mu(t), \quad z \in D,$$

where μ is a positive measure on Γ which is the weak star limit as $\varrho \to 1-$; of the measures $(1/2\pi) \log \|L_T(\varrho e^{it})\|\, dt$, $0 < \varrho < 1$.

Remark 1. As will be shown in Lemma 2.2, the assumption that R_T is in N_1 implies that $\sum_n (1-|\lambda_n|)<\infty$ and therefore (cf. [7, p. 54]) the above product converges uniformly on compact subsets of D to an inner H^∞ function. The existence of the limits which define the function w and the measure μ will be established in Proposition 3.1.

Remark 2. It is readily verified that the above formula for m_T can also be written in the form

$$m_T(z) = \lim_{\varrho\to 1-} z^k \prod_{\lambda_j\neq 0} \frac{\bar\lambda_j}{|\lambda_j|}\frac{\lambda_j - z}{1-\bar\lambda_j z} \exp \int_0^{2\pi} \frac{z+e^{it}}{z-e^{it}} \log \|\varrho+(1-\varrho)^2 e^{it} R_T(\varrho e^{it})\| \frac{dt}{2\pi}.$$

2. Preliminary results

In this section we present some preliminary results which are needed for the proof of Theorem 1.1. In the sequel, T will denote a fixed contraction in $\mathscr{L}(\mathscr{H})$. Following [12] we associate with T the self-adjoint operators

$$D_T = (I-T^*T)^{1/2} \quad\text{and}\quad D_{T^*} = (I-TT^*)^{1/2}$$

and set $\mathscr{D}_T=\overline{D_T\mathscr{H}}$ and $\mathscr{D}_{T^*}=\overline{D_{T^*}\mathscr{H}}$. In addition we denote by K_T the operator function defined by

$$K_T(\lambda) = D_T R_T(\lambda)(I-\lambda T^*), \quad \lambda\in\varrho(T),$$

and by U_T the set $\{x\in\mathscr{H}:\ \|D_{T^*}x\|\leqq 1\}$.

The first result of this section which will be needed in the proof of Theorem 1.1 appears in [12, p. 263], however it is expressed there in terms of the characteristic operator function, and one part of its proof depends on the functional model. In order to keep our exposition self-contained, we present below an equivalent formulation of this result, and give a proof which is similar to that in [12] but does not depend on the functional model and does not use explicitly the characteristic operator function.

Lemma 2.1. *For every λ in $D\cap\varrho(T)$*

$$\|L_T(\lambda)\| = \sup\{\|K_T(\lambda)x\|:\ x\in U_T\}.$$

Proof. We assume first that T is invertible, and prove the assertion for $\lambda=0$, that is, we show that

$$\|T^{-1}\| = \sup\{\|D_T T^{-1}x\|:\ x\in U_T\}.$$

Since $D_T T^{-1} = T^{-1} D_{T^*}$ [12, p. 7] this equality is equivalent to the equality $\|T^{-1}\| = \|J\|$ where J denotes the restriction of T^{-1} to $\mathscr{D}_{T^*}$. To show this, choose x in $\mathscr{H}$ and consider the orthogonal decomposition $x = x_1 + x_2$ where $x_1 \in \mathscr{D}_{T^*}$ and $x_2 \in \mathscr{D}_{T^*}^{\perp}$. Using the facts that T^{-1} maps $\mathscr{D}_{T^*}$ onto $\mathscr{D}_T$ and maps $\mathscr{D}_{T^*}^{\perp}$ isometrically onto $\mathscr{D}_T^{\perp}$ [12, p. 7], and that $\|J\| \geqq 1$ (since T is a contraction) we obtain that

$$\|T^{-1}x\|^2 = \|T^{-1}x_1\|^2 + \|T^{-1}x_2\|^2 \leqq \|J\|^2 \|x\|^2.$$

This shows that $\|T^{-1}\| \leqq \|J\|$, and since the reverse inequality is obvious, we conclude that $\|T^{-1}\| = \|J\|$. To prove the assertion in the general case, we assume that λ is in $D \cap \varrho(T)$, and consider the operator $T_\lambda = (\lambda I - T)(I - \bar{\lambda} T)^{-1}$, which is also a contraction [12, p. 14]. Since T_λ is invertible, we have by the assertion just proved that

$$\|L_T(\lambda)\| = \|T_\lambda^{-1}\| = \sup \{\|D_{T_\lambda} T_\lambda^{-1} x\| : x \in U_{T_\lambda}\}.$$

Setting $S = (1 - |\lambda|^2)^{1/2} (I - \bar{\lambda} T)^{-1}$, we obtain by a simple computation that

$$\|D_{T_\lambda} T_\lambda^{-1} x\|^2 = \langle D_{T_\lambda}^2 T_\lambda^{-1} x, T_\lambda^{-1} x \rangle = \|K_T(\lambda) S^* x\|^2$$

and noticing that $x \in U_{T_\lambda}$ if and only if $S^* x \in U_T$, we obtain the desired conclusion.

Lemma 2.2. *If F is a meromorphic function on D with values in some complex Banach space, with poles $\{\lambda_1, \lambda_2, \dots\}$ in D repeated according to multiplicity, then the following conditions are equivalent:*

(a) $\sup N(r, F) < \infty$,

(b) $\sum_n (1 - |\lambda_n|) < \infty$.

Proof. We assume that $n(0, F) = 0$. The general case can be reduced to this one by an obvious argument. We also assume that $|\lambda_n| \leqq |\lambda_{n+1}|$, $n = 1, 2, \dots$, and set $\nu(t) = n(t, F)$, $0 \leqq t < 1$. Integrating by parts and taking into account the assumption that $\nu(0) = 0$, we obtain that for every $0 \leqq r < 1$

$$\sum_{n=1}^{\nu(r)} \log r/|\lambda_n| = \int_0^r (\log r/t)\, d\nu(t) = \int_0^r (\nu(t)/t)\, dt = N(r, F).$$

This shows that condition (a) is equivalent to the condition $\sum_{n=1}^{\infty} \log 1/|\lambda_n| < \infty$ which is clearly equivalent to condition (b). This completes the proof.

From Lemma 2.2 we obtain an equivalent definition of the class $N_a(X)$:

Corollary 2.3. *If X is a complex Banach space and F is an X-valued meromorphic function on D, with poles $\{\lambda_1, \lambda_2, \dots\}$ repeated according to multiplicity,* *then F is in $N_\alpha(X)$ for some $\alpha \geqq 0$, if and only if* $\sup_{0 \leqq r<1} m_\alpha(r, F) < \infty$ *and* $\sum_n (1-|\lambda_n|) < \infty$.

Proof. This is an immediate consequence of Lemma 2.2 and the definition of the class $N_\alpha(X)$.

We recall that if $\{\lambda_1, \lambda_2, \ldots\}$ is a sequence in D which satisfies condition (b) of Lemma 2.2, and if k is the number of λ_n equal to zero, then the Blaschke product

$$B(z) = z^k \prod_{\lambda_n \neq 0} (\bar{\lambda}_n/|\lambda_n|)((\lambda_n - z)/(1-\bar{\lambda}_n z))$$

converges uniformly on compact subsets of D, and B is an inner function in H^∞ whose zeros in D are precisely the points λ_n, and each zero has multiplicity equal to the number of times it occurs in the sequence [7, p. 54].

If T is a contraction such that R_T is in N_α for some $\alpha \geqq 0$, then by Corollary 2.3 the sequence of poles of R_T in D (repeated according to multiplicity) satisfies condition (b) of Lemma 2.2, and therefore by the above observation the Blaschke product associated with this sequence is a well defined inner function in H^∞. We shall denote in the sequel this function by B_T.

Lemma 2.4. *If T is a contraction with resolvent in N_α for some $\alpha \geqq 0$, then the function* $\log \|B_T(z) L_T(z)\|$ *is subharmonic in D.*

Proof. Since the zeros of B_T coincide with the poles of R_T in D, including multiplicity, the operator function $B_T(z)K_T(z)$ is holomorphic in D, and therefore [7, p. 34], for every $x, y \in \mathscr{H}$, the function $\log |\langle B_T(z) K_T(z)x, y\rangle|$ is subharmonic in D. Hence, since by Lemma 2.1,

$$\log \|B_T(z) L_T(z)\| = \sup \{\log |\langle B_T(z) K_T(z) x, y\rangle| \colon x \in U_T,\ \|y\| \leqq 1\}$$

for all $z \in D \cap \varrho(T)$, it follows that the function $\log \|B_T(z) L_T(z)\|$ is subharmonic in $D \cap \varrho(T)$, and therefore since it is continuous in D and $D \setminus \varrho(T)$ is a discrete set, it follows by simple argument that it is also subharmonic in D.

We shall also need in the sequel the following elementary result which appears in [12, p. 263]. For the sake of completeness, we include the proof.

Lemma 2.5. *If T is a contraction then for every λ in $D \cap \varrho(T)$*

$$(1-|\lambda|)\|R_T(\lambda)\| \leqq \|L_T(\lambda)\| \leqq 1 + 2(1-|\lambda|)\|R_T(\lambda)\|.$$

Proof. Assume that λ is in $D \cap \varrho(T)$. Since $R_T(\lambda) = (I - \bar{\lambda} T)^{-1} L_T(\lambda)$ we have that $\|R_T(\lambda)\| \leqq \|(I-\bar{\lambda}T)^{-1}\| \|L_T(\lambda)\|$, and since T is a contraction,

$$\|(I-\bar{\lambda}T)^{-1}\| = \Big\|\sum_{n=0}^{\infty} \bar{\lambda}^n T^n\Big\| \leqq \sum_{n=0}^{\infty} |\lambda|^n = 1/(1-|\lambda|),$$

and consequently $(1-|\lambda|)\|R_T(\lambda)\| \leqq \|L_T(\lambda)\|$. The second inequality is an immediate consequence of the identity $L_T(\lambda)=\bar{\lambda}I+(1-|\lambda|^2)R_T(\lambda)$ which follows by a simple computation.

We shall also need in the sequel the fact that the class $C_{0\cdot}$ is invariant under certain Möbius transformation. This is given by:

Lemma 2.6. *If T is a contraction of class $C_{0\cdot}$ and $\alpha\in D$, then the operator $T_\alpha=(\alpha I-T)(I-\bar{\alpha}T)^{-1}$ is also of class $C_{0\cdot}$.*

Proof. Since T is a contraction T_α is also a contraction, and therefore the sequence of self-adjoint operators $T_\alpha^{*n}T_\alpha^n$ is decreasing, hence converges strongly to some self-adjoint operator L. The assertion that T_α is of class $C_{0\cdot}$ is clearly equivalent to the assertion that $L=0$. To prove this, notice that $T_\alpha^* L T_\alpha = L$, hence $(\bar{\alpha}I-T^*)L(\alpha I-T)=(I-\alpha T^*)L(I-\bar{\alpha}T)$ and therefore $T^*LT=L$. This implies that $T^{*n}LT^n=L$ for every positive integer n, so that for every $x\in\mathcal{H}$ we have that

$$\|Lx\| \leqq \|L\|\,\|T^n x\|, \quad n=0,1,2,\dots,$$

and consequently, since T is of class $C_{0\cdot}$, we conclude that $L=0$. This completes the proof.

3. Contractions with resolvent in N_1

We begin by showing that the limits in the definitions of the function w and the measure μ in the statement of Theorem 1.1, actually exist for every contraction with resolvent in N_1. This enables us to associate with every such contraction T a function φ_T in H^∞, which by virtue of Theorem 1.1, is a minimal function when T is of class C_0.

Proposition 3.1. *If T is a contraction with resolvent in N_1, then the measures $(1/2\pi)\log\|L_T(\varrho e^{it})\|\,dt$, $0\leqq\varrho<1$, converge as $\varrho\to 1-$, in the weak star topology to a positive measure μ on Γ. Furthermore, if w and φ_T are the holomorphic functions on D defined by*

$$w(z)=\int_\Gamma ((e^{it}+z)/(e^{it}-z))\,d\mu(t), \quad z\in D$$

and

$$\varphi_T(z)=B_T(z)\exp(-w(z)), \quad z\in D$$

then

$$w(z)=\lim_{\varrho\to 1-}(1/2\pi)\int_0^{2\pi}((e^{it}+z)/(e^{it}-z))\log\|L_T(\varrho e^{it})\|\,dt, \quad z\in D$$

and

$$\|\varphi_T(z)L_T(z)\| \leqq 1, \quad z\in D.$$

In addition, the function φ_T has the following minimality property: If f is a function in H^∞ which satisfies the condition

$$\|f(z)L_T(z)\| \leqq 1, \quad z\in D$$

then there exists a function h in H^∞ such that $\|h\|_\infty \leqq 1$ and $f=h\varphi_T$.

Proof. Since B_T is in H^∞ and $B_T \not\equiv 0$, we have that

$$\int_0^{2\pi} \left|\log |B_T(\varrho e^{i\theta})|\right| d\theta, \quad 0<\varrho<1,$$

is bounded as $\varrho\to 1$ (cf. [10, p. 90]) and therefore noticing that $\|L_T(\lambda)\|\geqq 1$, $\lambda\in D\cap\varrho(T)$ (since T is a contraction), we obtain from the assumption that R_T is in N_1 and Lemma 2.5, that

$$\int_0^{2\pi} \left|\log \|B_T(\varrho e^{i\theta})L_T(\varrho e^{i\theta})\|\right| d\theta, \quad 0<\varrho<1,$$

is also bounded as $\varrho\to 1$. Combining this with the fact that by Lemma 2.4 the function $\log\|B_T(z)L_T(z)\|$ is subharmonic in D, we infer (cf. [6] or [7, p. 38]) that the measures

$$(1/2\pi)\log \|B_T(\varrho e^{it})L_T(\varrho e^{it})\|\, dt, \quad 0<\varrho<1,$$

converge as $\varrho\to 1-$, in the weak star topology to a measure μ on Γ, and the function

$$u(z) = \int_\Gamma \operatorname{Re}\left((e^{it}+z)/(e^{it}-z)\right) d\mu(t), \quad z\in D$$

is the least harmonic majorant of the function $\log\|B_T(z)L_T(z)\|$ in D. Since B_T is a Blaschke product, we have that [7, p. 56]

$$\lim_{\varrho\to 1-} \int_0^{2\pi} \left|\log |B_T(\varrho e^{i\theta})|\right| d\theta = \lim_{\varrho\to 1-} \Big(-\int_0^{2\pi} \log |B_T(\varrho e^{i\theta})|\, d\theta\Big) = 0$$

and therefore μ is also the weak star limit as $\varrho\to 1-$, of the measures $(1/2\pi)\log\|L_T(\varrho e^{it})\|\, dt$, $0\leqq\varrho<1$. Thus remembering that $\|L_T(\lambda)\|\geqq 1$, $\lambda\in D\cap\varrho(T)$, we obtain that μ is a positive measure, and since for every $z\in D$ the function $e^{it}\to(e^{it}+z)/(e^{it}-z)$ is continuous on Γ, we also have that

$$w(z) = \lim_{\varrho\to 1-} (1/2\pi)\int_0^{2\pi} \left((e^{it}+z)/(e^{it}-z)\right)\log\|L_T(\varrho e^{it})\|\, dt, \quad z\in D.$$

It is also clear that $u(z)=\operatorname{Re} w(z)$, $z\in D$, and therefore from the above mentioned majorant property of u, we obtain that

$$\log\|B_T(z)L_T(z)\| \leqq \operatorname{Re} w(z), \quad z\in D$$

which is equivalent to the desired inequality,

$$\|\varphi_T(z)L_T(z)\| \leqq 1, \quad z\in D.$$

To prove the last assertion, assume that f is a function in H^∞ that satisfies the condition $\|f(z)L_T(z)\|\leqq 1$, $z\in D\cap\varrho(T)$. We may clearly assume that $f\not\equiv 0$. Since $\|L_T(z)\|\geqq 1$, $z\in D\cap\varrho(T)$, it follows by continuity that $|f(z)|\leqq 1$, $z\in D$. Consider the factorization $f=B\cdot g$, where B is the Blaschke product formed by the zeros of f in D. Then g is in H^∞ and $0<|g(z)|\leqq 1$, $z\in D$, [9, p. 66]. Using the hypothesis on f and Lemma 2.5 we obtain that

$$(1-|z|)\,|f(z)|\,\|R_T(z)\| \leqq 1, \quad z\in D\cap\varrho(T).$$

This implies that every pole of R_T in D, is a zero of f whose multiplicity is not less than the order of the pole. Thus $B=B_1\cdot B_T$, where B_1 is also a Blaschke product. Using again the hypothesis on f we obtain that

$$\log\|B(z)L_T(z)\| \leqq -\log|g(z)|, \quad z\in D\cap\varrho(T)$$

and by continuity this inequality also holds for all $z\in D$. Since $g(z)\neq 0$, $\forall z\in D$, the function $-\log|g(z)|$ is harmonic in D, hence is a harmonic majorant of the function $\log\|B(z)L_T(z)\|$ in D. But by Lemma 2.4 and the above factorization of B, this function is subharmonic in D, and therefore (by [7, p. 38] or [6]) its least harmonic majorant u_1 in D is given by

$$u_1(z) = \lim_{\varrho\to 1-} (1/2\pi)\int_0^{2\pi} \operatorname{Re}\left((e^{it}+z)/(e^{it}-z)\right)\log\|B(\varrho e^{it})L_T(\varrho e^{it})\|\,dt, \quad z\in D$$

and since B is a Blaschke product it follows by the argument already used in the proof of the first part of the proposition that

$$u_1(z) = \lim_{\varrho\to 1-} (1/2\pi)\int_0^{2\pi} \operatorname{Re}\left((e^{it}+z)/(e^{it}-z)\right)\log\|L_T(\varrho e^{it})\|\,dt = \operatorname{Re} w(z)$$

for all $z\in D$. Combining all these facts we obtain that

$$\operatorname{Re} w(z) \leqq -\log|g(z)|, \quad z\in D$$

hence the holomorphic function

$$h(z) = B_1(z)\,g(z)\exp\left(w(z)\right), \quad z\in D$$

satisfies the conditions $|h(z)|\leqq 1$, $z\in D$ and $f=h\varphi_T$. This concludes the proof of the proposition.

Remark. It is clear that $\lim\limits_{\varrho\to 1-}\|L_T(\varrho e^{it})\|=1$, uniformly on every compact subset of $\Gamma\setminus\sigma(T)$, and therefore the closed support of the measure μ defined in

Proposition 3.1 is contained in $\Gamma\cap\sigma(T)$. Hence in particular, if the set $\Gamma\cap\sigma(T)$ has linear measure zero, μ is a singular measure, and φ_T is an inner function.

We can now characterize contractions with resolvent in N_1.

Theorem 3.2. *If T is a contraction then the following conditions are equivalent:*

(a) *R_T is in N_1.*

(b) *R_T is a meromorphic operator function on D which admits a representation of the form*

$$R_T(\lambda)=G(\lambda)/\varphi(\lambda),\quad \lambda\in D\cap\varrho(T)$$

where φ is a function in H^∞ whose zeros in D coincide with the poles of R_T (including multiplicity) and G is a holomorphic operator function on D, which satisfies the condition

$$\sup_{\lambda\in D}(1-|\lambda|)\|G(\lambda)\|<\infty.$$

(c) *The set $D\cap\varrho(T)$ is not empty, and there exists a function $f\not\equiv 0$ in H^∞ such that*

$$\sup_{\lambda\in D\cap\varrho(T)}(1-|\lambda|)|f(\lambda)|\,\|R_T(\lambda)\|<\infty.$$

Proof. (a)$\Rightarrow$(b): If R_T is in N_1, then the zeros of the function φ_T (associated with T by Proposition 3.1) coincide with the poles of R_T, including multiplicity, and therefore the meromorphic function $\varphi_T R_T$ extends to a holomorphic operator function on D, which we denote by G. It follows from Proposition 3.1 and Lemma 2.5 that $(1-|\lambda|)\|G(\lambda)\|\leqq 1$ for $\lambda\in D\cap\varrho(T)$, and by continuity this inequality holds also for all $\lambda\in D$. Hence condition (b) is satisfied with $\varphi=\varphi_T$ and $G=\varphi_T R_T$.

(b)$\Rightarrow$(c): This is obvious.

(c)$\Rightarrow$(a): Assume that condition (c) holds for some function $f\not\equiv 0$ in H^∞, and denote for every $\lambda\in\varrho(T)$ by $d(\lambda)$ the distance of λ from $\sigma(T)$. Since for every $\lambda\in\varrho(T)$ we have the inequality $(d(\lambda))^{-1}\leqq\|R_T(\lambda)\|$, (cf. [5, p. 567]) it follows from the assumption on f that for some constant $M>0$

$$|f(\lambda)|\leqq Md(\lambda)/(1-|\lambda|),\quad \lambda\in D\cap\varrho(T)$$

and therefore by continuity, f vanishes on $D\cap\sigma(T)$. Consequently, since f is holomorphic and $f\not\equiv 0$, the set $D\cap\sigma(T)$ is discrete, and therefore by condition (*c*), all the singularities of R_T in D are poles, and the order of each pole does not exceed its multiplicity as a zero of f. Thus R_T is meromorphic on D, and by the Blaschke condition satisfied by the zeros of a function in H^∞ [7, p. 53], we obtain that the sequence of poles of R_T in D satisfies condition (b) of Lemma 2.2. Condition (c) also implies that there exists a constant $K>0$ such that

$$\int_0^{2\pi} \log^+ \|(1-r)R_T(re^{i\theta})\|\, d\theta \leqq \int_0^{2\pi} |\log |f(re^{i\theta})||\, d\theta + K$$

for all $0\leqq r<1$. Since $f\in H^\infty$ and $f\not\equiv 0$, the expression on the right hand side of the above inequality is dominated by a positive constant which does not depend on r [10, p. 90]. Thus by Corollary 2.3 we conclude that R_T is in N_1, and the proof of the theorem is complete.

We conclude this section with a result that describes the action of certain Möbius transformations on contractions with resolvent in N_1. This result will be required for the proof of Theorem 1.1.

Proposition 3.3. *Let T be a contraction with resolvent in N_1. Fix $\alpha\in D$ and consider the function $q(z)=(\alpha-z)/(1-\bar{\alpha}z)$, $z\in D$. Then the contraction $T_\alpha=q(T)$ has also resolvent in N_1, and there exists a constant c of modulus* 1, *such that*

$$\varphi_{T_\alpha}(z) = c\varphi_T(q(z)), \quad z\in D.$$

Proof. A simple computation shows that $\lambda\in\varrho(T_\alpha)$ if and only if $q(\lambda)\in\varrho(T)$ and

$$R_{T_\alpha}(\lambda) = (\bar{\alpha}\lambda-1)^{-1}(I-\bar{\alpha}T)R_T(q(\lambda)), \quad \lambda\in\varrho(T_\alpha).$$

Thus, using the representation of R_T given by part (b) of Theorem 3.2, we obtain that

$$R_{T_\alpha}(\lambda) = G_1(\lambda)/\varphi_1(\lambda), \quad \lambda\in D\cap\varrho(T_\alpha)$$

where $\varphi_1(\lambda)=(\bar{\alpha}\lambda-1)^{-1}\varphi(q(\lambda))$ and $G_1(\lambda)=(1-\bar{\alpha}T)G(q(\lambda))$ for every $\lambda\in D$. Hence remembering that

$$\sup_{\lambda\in D}(1-|\lambda|)\|G(\lambda)\| < \infty$$

and using the estimate

$$(1-|\lambda|)/(1-|q(\lambda)|) \leqq 4/(1-|\alpha|), \quad \lambda\in D$$

(see [7, p. 3, formula 1.5]) we obtain that

$$\sup_{\lambda\in D}(1-|\lambda|)\|G_1(\lambda)\| < \infty$$

and therefore by Theorem 3.2, R_{T_α} is in N_1.

We turn now to the proof of the second assertion. A direct computation shows that for every $\lambda\in D\cap\varrho(T_\alpha)$

$$L_{T_\alpha}(\lambda) = (1-\alpha\bar{\lambda})(\bar{\alpha}\lambda-1)^{-1}L_T(q(\lambda))$$

and therefore $\|L_{T_\alpha}(\lambda)\| = \|L_T(q(\lambda))\|$. Hence using the fact that by Proposition 3.1,

$$\|\varphi_T(\lambda)L_T(\lambda)\| \leqq 1, \quad \lambda\in D\cap\varrho(T)$$

we obtain that also

$$\|\varphi_T(q(\lambda)L_{T_\alpha}(\lambda)\| \leqq 1, \quad \lambda\in D\cap\varrho(T_\alpha)$$

and therefore by the minimality property of the function φ_{T_α}, there exists a function g in H^∞, such that $\|g\|_\infty \leqq 1$ and

$$\varphi_{T_\alpha}(\lambda) = g(\lambda)\varphi_T(q(\lambda)), \quad \lambda\in D.$$

Changing the roles of T_α and T and noticing that $q(q(\lambda))=\lambda$, $\forall\lambda\in D$, we obtain in the same way, that there exists a function h in H^∞, such that $\|h\|_\infty \leqq 1$ and

$$\varphi_T(\lambda) = h(\lambda)\varphi_{T_\alpha}(q(\lambda)), \quad \lambda\in D$$

and therefore

$$\varphi_T(q(\lambda)) = h(q(\lambda))\varphi_{T_\alpha}(\lambda), \quad \lambda\in D.$$

Hence by the maximum principle $g\equiv c$, where c is a constant of modulus 1. This concludes the proof.

4. Proof of the main result

For the proof of Theorem 1.1 we require one more preliminary result.

Lemma 4.1. *Let T be a completely non unitary contraction such that R_T is in N_1. Then setting $\varphi=\varphi_T$, we have that*

$$(1-|\lambda|)\|\varphi(T)R_T(\lambda)\| \leqq 3, \quad \lambda\in D\cap\varrho(T).$$

Proof. For every $\lambda\in D$ consider the holomorphic function h_λ on D defined by

$$h_\lambda(z) = (\varphi(z)-\varphi(\lambda))(z-\lambda)^{-1}, \quad z\in D, \quad z\neq\lambda.$$

It is easily verified that $h_\lambda\in H^\infty$ and $\|h_\lambda\|_\infty \leqq 2/(1-|\lambda|)$, and therefore also $\|h_\lambda(T)\| \leqq$ $\leqq 2/(1-|\lambda|)$. Since

$$\varphi(T)-\varphi(\lambda)I = (T-\lambda I)h_\lambda(T), \quad \lambda\in D$$

it follows that

$$(\varphi(T)-\varphi(\lambda)I)R_T(\lambda) = -h_\lambda(T), \quad \lambda\in D\cap\varrho(T)$$

and consequently

$$\|(\varphi(T)-\varphi(\lambda)I)R_T(\lambda)\| \leqq 2/(1-|\lambda|), \quad \lambda\in D\cap\varrho(T).$$

This implies the desired conclusion by virtue of Lemma 2.5 and Proposition 3.1.

Proof of Theorem 1.1. We assume first that T is an invertible contraction of class C_0 with resolvent in N_1, and prove that T is annihilated by φ_T. For this, we set $\varphi=\varphi_T$, and consider the operator function

$$F(\lambda) = T\varphi(T)R_T(\lambda), \quad \lambda\in\varrho(T).$$

Observe that since T is of class $C_{0\cdot}$, it is completely non unitary, and therefore $\varphi(T)$ is well defined. The singularities of F in D, which are the poles of R_T, are removable, since by Lemma 4.1 we have for every $0<r<1$; that

$$\sup\{\|F(\lambda)\|: \lambda\in\varrho(T),\ |\lambda|<r\}<\infty.$$

Thus F is holomorphic in D, and therefore using the assumption that T is invertible we obtain from the Taylor expansion of R_T around $z=0$, that

$$F(\lambda) = -\sum_{n=0}^{\infty}\varphi(T)T^{-n}\lambda^n, \quad \lambda\in D$$

the series converging in the operator norm. Combining this with the Laurent expansion of R_T for $|z|>1$, we obtain that for every $re^{i\theta}\in D$,

$$F(re^{i\theta})-F(r^{-1}e^{i\theta}) = -\sum_{n=-\infty}^{\infty} r^{|n|}\varphi(T)T^{-n}e^{in\theta}$$

the series converging again in the operator norm. On the other hand, using the resolvent identity

$$R_T(\lambda)-R_T(\lambda') = (\lambda'-\lambda)R_T(\lambda)R_T(\lambda'), \quad \lambda,\lambda'\in\varrho(T)$$

we obtain that for every $re^{i\theta}\in D$

$$F(re^{i\theta})-F(r^{-1}e^{i\theta}) = e^{i\theta}r^{-1}(1-r^2)F(re^{i\theta})R_T(r^{-1}e^{i\theta})$$

and therefore by Lemma 4.1, we obtain that for every $x\in\mathscr{H}$ and $re^{i\theta}\in D$

$$\|(F(re^{i\theta})-F(r^{-1}e^{i\theta}))x\| \leqq 6r^{-1}\|R_T(r^{-1}e^{i\theta})x\|.$$

Hence applying the Parseval identity for Hilbert space valued functions on Γ, we obtain that for every $x\in\mathscr{H}$

$$\sum_{n=-\infty}^{\infty} r^{2|n|}\|\varphi(T)T^{-n}x\|^2 = (1/2\pi)\int_0^{2\pi}\|(F(re^{i\theta})-F(r^{-1}e^{i\theta}))x\|^2\,d\theta \leqq$$

$$\leqq 36r^{-2}(1/2\pi)\int_0^{2\pi}\|R_T(r^{-1}e^{i\theta})x\|^2\,d\theta = 36\sum_{n=0}^{\infty}r^{2n}\|T^nx\|^2.$$

(The proof of this inequality was inspired by the methods in [16].) Since T is a contraction

$$\|T^{-n}\varphi(T)x\| \leqq \|T^{-n-1}\varphi(T)x\|, \quad n=0,1,2,\ldots,$$

and therefore

$$\|\varphi(T)x\|^2 \leqq (1-r^2) \sum_{n=-\infty}^{\infty} r^{2|n|} \|T^{-n}\varphi(T)x\|^2.$$

Combining this with the preceeding estimate we obtain that

$$\|\varphi(T)x\|^2 \leqq 36(1-r^2) \sum_{n=0}^{\infty} r^{2n} \|T^n x\|^2.$$

But the assumption that T is of class $C_{0.}$ implies that the expression on the right hand side of the above inequality tends to zero as $r \to 1-$, and consequently, $\varphi(T)x=0$. Since this holds for every $x \in \mathscr{H}$, we conclude that $\varphi(T)=0$. To prove the same result for T not necessarily invertible, assume again that T is of class $C_{0.}$ and that R_T is in N_1. Choose $\alpha \in D \cap \varrho(T)$, and consider the function $q(z)= =(\alpha-z)/(1-\bar{\alpha}z)$, $z \in D$, and the invertible contraction $T_\alpha = q(T)$. By Lemma 2.6 and Proposition 3.3, T_α is also of class $C_{0.}$ and has resolvent in N_1, and therefore by what has just been proved, we have that $\varphi_\alpha(T)=0$ where φ_α denotes the function φ_{T_α}. But by Proposition 3.3, $\varphi_\alpha = c\varphi \circ q$ where c is a constant of modulus 1, and therefore using the fact that $q \circ q(\lambda)=\lambda$, $\forall \lambda \in D$, we obtain that $\varphi = c^{-1}\varphi_\alpha \circ q$ and consequently $\varphi(T)=c^{-1}\varphi_\alpha(T_\alpha)=0$. This establishes the assertion in the general case.

We show next that φ_T is a minimal function of T. For this assume that f is a function in H^∞ such that $\|f\|_\infty \leqq 1$ and $f(T)=0$. To prove that φ_T divides f in H^∞, consider for every $\lambda \in D$, the holomorphic function g_λ on D defined by

$$g_\lambda(z) = (f(z)-f(\lambda))(1-\overline{f(\lambda)}f(z))^{-1}(z-\lambda)^{-1}(1-\bar{\lambda}z), \quad z \in D, \quad z \neq \lambda.$$

Since $\|f\|_\infty \leqq 1$ also $\|g_\lambda\|_\infty \leqq 1$ and therefore also $\|g_\lambda(T)\| \leqq 1$. Using the identity

$$g_\lambda(z)(1-\overline{f(\lambda)}f(z)) = (f(z)-f(\lambda))(1-\bar{\lambda}z)(z-\lambda)^{-1}$$

and the assumption that $f(T)=0$, we obtain that for every $\lambda \in D \cap \varrho(T)$,

$$\|f(\lambda)L_T(\lambda)\| = \|g_\lambda(T)\| \leqq 1.$$

Consequently, by Proposition 3.1, there exists a function $h \in H^\infty$ such that $\|h\|_\infty \leqq 1$ and $f=h\varphi_T$. Hence φ_T divides every function in H^∞ which annihilates T. We show now that φ_T is an inner function. For this, set again $\varphi=\varphi_T$ and consider the canonical factorization $\varphi=\varphi_1 \cdot \varphi_2$ where φ_1 and φ_2 are the outer and inner factors of φ, respectively. Then $\varphi_1(T)\varphi_2(T)=0$, and therefore using the fact that $\varphi_1(T)$ has zero kernel since φ_1 is outer [12, p. 118], we obtain that also $\varphi_2(T)=0$. Hence by the result just proved, we have that

$$|\varphi_1(\lambda)|^{-1} = |\varphi_2(\lambda) \cdot \varphi(\lambda)^{-1}| \leqq 1, \quad \text{for all} \quad \lambda \in D.$$

On the other hand since $\|\varphi\|_\infty \leqq 1$ also $\|\varphi_1\|_\infty \leqq 1$, and therefore by the maximum principle $\varphi_1 \equiv c$, where c is a constant of modulus 1. Thus φ_T is an inner function, and consequently is a minimal function of T. (Since $\varphi(0)$ is real we actually have that $c=1$, so that $\varphi=\varphi_2$.)

Finally to prove the remaining assertion of the theorem, assume that T is a contraction of class C_0. Then by [12, p. 123] T is of class $C_{0\cdot}$. (For a proof of this fact which is independent of dilation theory see [8].) To show that R_T is in N_1, consider a function $f \not\equiv 0$ in H^∞ such that $f(T)=0$. For every $\lambda \in D$, consider the holomorphic function f_λ on D defined by

$$f_\lambda(z) = \big(f(z)-f(\lambda)\big)(z-\lambda)^{-1}, \quad z \in D, \quad z \neq \lambda.$$

Then $\|f_\lambda\|_\infty \leqq 2\|f\|_\infty/(1-|\lambda|)$ and therefore also $\|f_\lambda(T)\| \leqq 2\|f\|_\infty/(1-|\lambda|)$. Since $f(T)=0$ we have that $f_\lambda(T)(\lambda I-T)=f(\lambda)I$, hence if $f(\lambda)\neq 0$ then $\lambda \in D \cap \varrho(T)$, and therefore since $f \not\equiv 0$, the set $D \cap \varrho(T)$ is not empty. It also follows from the preceeding facts that

$$\|f(\lambda)R_T(\lambda)\| = \|f_\lambda(T)\| \leqq 2\|f\|_\infty/(1-|\lambda|),$$

for all λ in $D \cap \varrho(T)$. Thus R_T satisfies condition (c) of Theorem 3.2 and therefore by that theorem R_T is in N_1. This concludes the proof of Theorem 1.1.

Remark 1. Observe that we obtained above also a proof of the existence of a minimal function for a contraction of class C_0 which is different from the proof of this fact given in [12, p. 124]. Still another proof of this fact appears in [3, p. 188].

Remark 2. It follows from Proposition 3.1 and the proof of Theorem 1.1, that if T is a contraction of class C_0 then a function f in H^∞ annihilates T, if and only if, it satisfies the condition

$$\sup_{\lambda \in D \cap \varrho(T)} |f(\lambda)|\,\|L_T(\lambda)\| < \infty$$

which by virtue of Lemma 2.5 is also equivalent to the condition

$$\sup_{\lambda \in D \cap \varrho(T)} (1-|\lambda|)\,|f(\lambda)|\,\|R_T(\lambda)\| < \infty.$$

An immediate consequence of Theorem 1.1 and Theorem 3.2 is the following:

Corollary 4.2 (Dazord [4]). *A contraction T is of class C_0, if and only if, T is of class $C_{0\cdot}$, and there exists a function $f \not\equiv 0$ in H^∞ such that*

$$\sup_{\lambda \in D \cap \varrho(T)} (1-|\lambda|)\,|f(\lambda)|\,\|R_T(\lambda)\| < \infty.$$

5. Contractions whose characteristic function has a scalar multiple

In this section we express the preceeding results in terms of the characteristic operator function associated with a contraction. We recall that [12, Chapter VI] if T is a contraction then its characteristic function is the holomorphic operator function θ_T on D, whose value at every $\lambda \in D$ is the bounded linear operator $\theta_T(\lambda)$ from the Hilbert space $\mathscr{D}_T$ to the Hilbert space $\mathscr{D}_{T^*}$, which is defined by the relation

$$\theta_T(\lambda) D_T = D_{T^*}(I-\lambda T^*)^{-1}(\lambda I-T).$$

For every $\lambda \in \varrho(T)$ the operator $\theta_T(\lambda)$ is invertible and its inverse $\theta_T(\lambda)^{-1}$ is the bounded linear operator from $\mathscr{D}_{T^*}$ to $\mathscr{D}_T$ which satisfies the equality

$$\theta_T(\lambda)^{-1} D_{T^*} = D_T(\lambda I-T)^{-1}(I-\lambda T^*) = K_T(\lambda)$$

(K_T is the operator function defined in Section 2). Thus from Lemma 2.1 we obtain for every $\lambda \in \varrho(T)$ the equality $\|L_T(\lambda)\| = \|\theta_T(\lambda)^{-1}\|$, which is also proved in [12, p. 264]. This implies by Lemma 2.5, that R_T is meromorphic in D if and only if θ_T^{-1} is meromorphic in D, and in this case these two functions have the same poles in D with the same orders (see also [12, p. 264]). Thus using Lemma 2.5 we obtain:

Proposition 5.1. *If T is a contraction then R_T is in N_1, if and only if, θ_T^{-1} is of bounded characteristic (as a function from $D \cap \varrho(T)$ into the Banach space of all bounded linear operators from $\mathscr{D}_{T^*}$ to $\mathscr{D}_T$).*

We recall [12, p. 264] that a function $f \not\equiv 0$ in H^∞ is called a scalar multiple of θ_T, if there exists a holomorphic operator function Ω on D whose values are bounded linear operators from $\mathscr{D}_{T^*}$ to $\mathscr{D}_T$ with norm not exceeding 1, such that for every $\lambda \in D$

$$\Omega(\lambda)\theta_T(\lambda) = f(\lambda) I_1 \quad \text{and} \quad \theta_T(\lambda)\Omega(\lambda) = f(\lambda) I_2$$

where I_1 and I_2 are the identity operators on $\mathscr{D}_T$ and $\mathscr{D}_{T^*}$ respectively.

The above equalities are clearly equivalent to the equality

$$\theta_T(\lambda)^{-1} = \Omega(\lambda)/f(\lambda)$$

for every $\lambda \in D$ such that $f(\lambda) \neq 0$. Thus from [2, Th. 2.1] and Proposition 5.1 we obtain

Theorem 5.2. *If T is a contraction then the following conditions are equivalent:*
(a) *R_T is in N_1,*
(b) *θ_T^{-1} is of bounded characteristic,*
(c) *θ_T has a scalar multiple.*

Remark. It follows from Theorem 5.2, Proposition 3.1 and the identity $\|\theta_T(\lambda)^{-1}\|=\|L_T(\lambda)\|$, $\lambda\in D\cap\varrho(T)$, that if θ_T has a scalar multiple f, then φ_T is also a scalar multiple of θ_T, and there exists a function h in H^∞ such that $\|h\|_\infty\leqq 1$ and $f=h\varphi_T$. Thus φ_T divides every other scalar multiple of θ_T and therefore can be called a minimal scalar multiple of θ_T.

Finally, from Theorem 1.1, Theorem 5.2 and the equality $\|\theta_T(\lambda)^{-1}\|=\|L_T(\lambda)\|$, $\lambda\in D\cap\varrho(T)$, we obtain a characterization of operators of class C_0 and a formula for their minimal function expressed in terms of the characteristic function:

Theorem 5.3. *A contraction T is of class C_0 if and only if T is of class $C_{0\cdot}$ and one of the three equivalent conditions of Theorem 5.2 is satisfied. Furthermore, if T is of class C_0 it has a minimal function given by*

$$m_T(z)=B_T(z)\exp\left(-w(z)\right),\quad z\in D$$

where

$$w(z)=\lim_{\varrho\to 1-}(1/2\pi)\int_0^{2\pi}\left((e^{it}+z)/(e^{it}-z)\right)\log\|\theta_T(\varrho e^{it})^{-1}\|\,dt,\quad z\in D$$

or alternatively

$$w(z)=\int_\Gamma\left((e^{it}+z)/(e^{it}-z)\right)d\mu(t),\quad z\in D$$

where

$$\mu=w^*\lim_{\varrho\to 1-}(1/2\pi)\log\|\theta_T(\varrho e^{it})^{-1}\|\,dt.$$

An immediate consequence of Theorem 5.3 is

Corollary 5.4 (Sz.-Nagy and Foiaş [12, p. 265]). *A contraction T is of class C_0 if and only if T is of class $C_{0\cdot}$ and θ_T has a scalar multiple.*

6. Contractions which are annihilated by functions in the disc algebra

Let A denote the disc algebra, that is the Banach algebra of all continuous functions on the closed unit disc $\bar{D}$ which are holomorphic in D, equipped with the supremum norm. In view of the von Neumann inequality [12, p. 32] and the fact that the polynomials are dense in A, there exists for every contraction T a norm continuous multiplicative homomorphism of the Banach algebra A into the Banach algebra $\mathscr{L}(\mathscr{H})$, which extends the mapping $p\to p(T)$ where p is a polynomial (see also [3, p. 167] and [8]). It is easily verified that if $f(z)=\sum_{n=0}^\infty a_nz^n$ is a function in A, then the operator $f(T)$ which corresponds to f by this homomorphism, is given by $f(T)=\lim_{r\to 1-}\sum_{n=0}^\infty a_nr^nT^n$ (where the convergence of the series and the limit are in the operator norm). It is also clear that if T is completely non unitary, then this homomorphism is the restriction to A of the homomorphism from H^∞ into $\mathcal{L}(\mathcal{H})$ given by [12, p. 117, Theorem 2.3].

Following [8] we shall say that a contraction T is of class D_0 if there exists a function $f\not\equiv 0$ in A such that $f(T)=0$.

The characterization of contractions of class D_0 is given by:

Theorem 6.1. *A contraction T is of class D_0, if and only if R_T is in N_1 and the set $\sigma(T)\cap\Gamma$ has linear measure zero.*

Proof. We assume first that T is a contraction such that R_T is in N_1 and $\sigma(T)\cap\Gamma$ has linear measure zero, and show that T is of class D_0. Since the set $\sigma(T)\cap\Gamma$ has linear measure zero, there exists by a Theorem of Fatou [9, p. 80] a function $g\not\equiv 0$ in A such that $g=0$ on $\sigma(T)\cap\Gamma$. (One can also choose by that theorem, g to be outer and so that it vanishes only on $\sigma(T)\cap\Gamma$.) Consider the function $f=g\varphi_T$. We claim that f is in A and that $f(T)=0$. By the remark following the proof of Propositon 3.1, the closed support of the measure μ which is associated with φ_T, is contained in $\sigma(T)\cap\Gamma$, and therefore since this set has linear measure zero (by assumption) μ is a singular measure and φ_T is an inner function. Since the accumulation points of the zeros of B_T are also contained in $\sigma(T)\cap\Gamma$; it follows [9, p. 68] that φ_T is continuous on $\bar{D}\setminus(\sigma(T)\cap\Gamma)$. Therefore, since $g=0$ on $\sigma(T)\cap\Gamma$, the function f extends to a continuous function on $\bar{D}$ which vanishes on $\sigma(T)\cap\Gamma$. Thus f is in A.

To show that $f(T)=0$, consider the canonical decomposition $T=T_0\oplus T_1$ [12, p. 9] where T_0 and T_1 are the unitary and completely non unitary parts of T respectively. Since $f(T)=f(T_0)\oplus f(T_1)$, we have to show that $f(T_0)=0$ and $f(T_1)=0$. The fact that T_0 is unitary implies that $\sigma(T_0)\subset\sigma(T)\cap\Gamma$, and therefore, since $f=0$ on $\sigma(T)\cap\Gamma$, it follows from the spectral theorem for unitary operators that $f(T_0)=0$. To show that also $f(T_1)=0$, observe first that for every $\lambda\in\varrho(T)$

$$R_T(\lambda) = R_{T_0}(\lambda)\oplus R_{T_1}(\lambda) \quad\text{and}\quad \sigma(T)\cap D = \sigma(T_1)\cap D.$$

This implies that R_{T_1} is also in N_1 and $B_T=B_{T_1}$. Also, using the facts that for every $\lambda\in\varrho(T)$, $L_T(\lambda)=L_{T_0}(\lambda)\oplus L_{T_1}(\lambda)$ and $\|L_{T_0}(\lambda)\|=1$ (since T_0 is unitary) and remembering that $\|L_{T_1}(\lambda)\|\geqq 1$; $\lambda\in\varrho(T_1)\cap D$, we obtain that

$$\|L_T(\lambda)\| = \|L_{T_1}(\lambda)\|,\quad \lambda\in\varrho(T)\cap D = \varrho(T_1)\cap D,$$

and combining this with the equality $B_T=B_{T_1}$, we infer that $\varphi_T=\varphi_{T_1}$. Since the set $\sigma(T)\cap\Gamma$ has linear measure zero, the same is true for its subset $\sigma(T_1)\cap\Gamma$, and therefore, since T_1 is completely non unitary, it follows [12, p. 84, Proposition 6.7] that T_1 is of class $C_{0\cdot}$. Thus by Theorem 1.1, T_1 is annihilated by φ_{T_1} and since $\varphi_{T_1}=\varphi_T$ we also have that $f(T_1)=0$. This proves that T is of class D_0.

To prove the converse assume that T is a contraction of class D_0. Then by [1, Corollary 4] or [8], the set $\sigma(T)\cap\Gamma$ has linear measure zero. The proof that R_T is in N_1 is exactly the same as the proof of the last part of Theorem 1.1. The only additional fact to observe is, that if f is a function in A, then for every $\lambda\in D$ the holomorphic function on D defined by

$$f_\lambda(z) = \big(f(z)-f(\lambda)\big)(z-\lambda)^{-1}, \quad z\in\bar{D}, \quad z\neq\lambda$$

is also in A. This concludes the proof of the theorem.

The preceeding proof shows that if T is a contraction of class D_0, then for every function g in A that vanishes on $\sigma(T)\cap\Gamma$, the function $f=g\varphi_T$ is also in A and $f(T)=0$. We show next that this is the general form of a function in A which annihilates T.

Proposition 6.2. *If T is a contraction of class D_0 and f is a function in A, then $f(T)=0$, if and only if there exists a function g in A that vanishes on $\sigma(T)\cap\Gamma$ such that $f=g\varphi_T$.*

Proof. In view of the preceeding observation it remains to show that every function in A which annihilates T, is of the above form. To show this assume that f is a function in A such that $f(T)=0$. Then also $f(T_1)=0$; where T_1 denotes as before the completely non unitary part of T. Therefore by Theorem 1.1, there exists a function g in H^∞ such that $f=g\varphi_{T_1}$, and since by the proof of Theorem 6.1, $\varphi_{T_1}=\varphi_T$, we have that $f=g\varphi_T$. Since φ_T is an inner function which (as observed in the proof of Theorem 6.1) is continuous on $\bar{D}\setminus(\sigma(T)\cap\Gamma)$; and since as shown in the proof of [1, Th. 4], $f=0$ on $\sigma(T)$, it follows that g extends to a continuous function on $\bar{D}$, which vanishes on $\sigma(T)\cap\Gamma$. This completes the proof of the proposition.

Remark. As observed in [8], it follows from the characterization of closed ideals in the algebra A [9, p. 85] that every contraction T of class D_0 determines uniquely a closed set $K\subset\Gamma$ of linear measure zero, and an inner function φ, such that a function f in A annihilates T, if and only if $f=g\varphi$, where g is a function in A that vanishes on K. Proposition 6.2 gives an independent proof of this fact and also provides the more precise information that $K=\sigma(T)\cap\Gamma$ and $\varphi=\varphi_T$.

7. Contractions with resolvent in N_α for some $0\leqq\alpha<1$

According to [2, Theorem 1.2] a contraction T has resolvent of bounded characteristic, if and only if, T is of class D_0 and $\sigma(T)$ is a thin set, that is, in addition to the Blaschke condition satisfied by the countable set $\sigma(T)\cap D$, also the condition

$$\int_0^{2\pi} (\log 1/d(e^{i\theta}, \sigma(T)))\, d\theta < \infty$$

holds (where for $\lambda \in \mathbf{C}$, $d(\lambda, \sigma(T))$ denotes the distance of λ from $\sigma(T)$).

For contractions with resolvent in N_α for some $0<\alpha<1$, we only have a partial result:

Theorem 7.1. *Assume that T is a contraction such that R_T is in N_α for some $0<\alpha<1$. Then T is of class D_0 and*

$$\int_0^{2\pi} (\log^+ 1/d(e^{i\theta}, \sigma(T)))^{1-\delta}\, d\theta < \infty$$

for every $\delta>0$.

Proof. We prove first the second assertion of the theorem. To simplify notations we set $d(\lambda)=d(\lambda, \sigma(T))$, for every $\lambda \in \mathbf{C}$. Remembering that [5, p. 567]

$$(d(\lambda))^{-1} \leqq \|R_T(\lambda)\|, \quad \lambda \in \varrho(T)$$

and using the assumption that R_T is in N_α, we obtain that there exists a constant $M>0$ such that

$$\int_0^{2\pi} (\log (1-r)^\alpha / d(re^{i\theta}))\, d\theta \leqq M, \quad 0 \leqq r < 1.$$

For every $t>0$, consider the set

$$E_t = \{\theta \in [0, 2\pi)\colon\ d(e^{i\theta}) \leqq t\}$$

and denote its Lebesgue measure by $m(t)$. Thus m is the distribution function of the function $\theta \to d(e^{i\theta})$, $\theta \in [0, 2\pi)$. Noticing that

$$d(re^{i\theta}) \leqq (1-r) + d(e^{i\theta}), \quad e^{i\theta} \in \Gamma, \quad 0 \leqq r < 1,$$

we obtain from the preceeding inequality that for every $0<t\leqq 1$,

$$m(t) \log 1/2t^{1-\alpha} \leqq \int_{E_t} (\log t^\alpha / d((1-t)e^{i\theta}))\, d\theta \leqq M$$

and therefore since $\alpha<1$ we deduce that there exists a positive constant c such that

$$m(t) \leqq c(2+\log 1/t)^{-1}, \quad 0 < t < 2\pi.$$

It is also clear that $m(t)=2\pi$ for $t \geqq 2\pi$. Thus using the well known properties of the distribution function (cf. [10, p. 65]) we obtain by integrating by parts and using the estimate above, that for every $\delta > 0$

$$\int_0^{2\pi} (\log^+ 1/d(e^{i\theta}))^{1-\delta}\, d\theta \leqq \int_0^{2\pi} (2+\log 1/d(e^{i\theta}))^{1-\delta}\, d\theta =$$

$$= \int_0^{2\pi} (2+\log 1/t)^{1-\delta}\, dm(t) \leqq 2\pi + c(1-\delta)\int_0^{2\pi} (2+\log 1/t)^{-1-\delta}(1/t)\, dt.$$

Since $\delta > 0$, the last integral converges, and the assertion is established.

To prove that T is of class D_0, denote for every $e^{i\theta}\in\Gamma$ by $d_1(e^{i\theta})$ the distance of $e^{i\theta}$ from the set $\Gamma\cap\sigma(T)$, and fix $0<\delta<1$. Then by the assertion just proved,

$$\int_0^{2\pi} (\log^+ 1/d_1(e^{i\theta}))^{1-\delta}\, d\theta < \infty$$

and this clearly implies that the set $\Gamma\cap\sigma(T)$ has linear measure zero. Thus by Theorem 6.1, T is of class D_0.

References

[1] A. Atzmon, Operators which are annihilated by analytic functions and invariant subspaces, *Acta Math.*, **144** (1980), 27—63.

[2] A. Atzmon, Operators with resolvent of bounded characteristic, *J. Integral Equations and Operator Theory*, **6** (1983), 779—803.

[3] R. G. Douglas, Canonical models, in: *Topics in operator theory*, Mathematical Surveys, no. 13, Amer. Math. Soc. (Providence, R. I., 1974); pp. 163—218.

[4] J. Dazord, The growth of the resolvent of a contraction of class C_0, *Rev. Roumaine Math. Pures Appl.*, **24** (1979), 213—218.

[5] N. Dunford and J. T. Schwartz, *Linear operators*, part I, Interscience (New York, 1968).

[6] L. Gårdind and L. Hörmander, Strongly subharmonic functions, *Math. Scand.*, **13** (1974), 93—96.

[7] J. B. Garnett, *Bounded analytic functions*, Academic Press (New York, 1981).

[8] P. Ghatage, Generalized algebraic operators, *Proc. Amer. Math. Soc.*, **52** (1975), 232—236.

[9] K. Hoffman, *Banach spaces of analytic functions*, Prentice Hall (Englewood Cliffs, N. J., 1962).

[10] Y. Katznelson, *An introduction to harmonic analysis*, Dover (New York, 1976).

[11] B. Sz.-Nagy and C. Foiaş, Sur les contractions de l'espace de Hilbert. VII, Triangulations canoniques. Fonction minimum, *Acta Sci. Math.*, **25** (1964), 12—37.

[12] B. Sz.-Nagy and C. Foiaş, *Harmonic analysis of operators on Hilbert space*, North Holland (Amsterdam, 1970).

[13] B. Sz.-Nagy and C. Foiaş, Modéle de Jordan pour une classe d'opérateurs de l'espace de Hilbert, *Acta Sci. Math.*, **31** (1970), 91—115.

[14] B. Sz.-Nagy and C. Foiaş, Compléments à l'étude des opérateurs de classe C_0, *Acta Sci. Math.*, **31** (1970), 287—296, and **33** (1972), 113—116.

[15] R. Nevanlinna, *Analytic functions*, Springer-Verlag (New York, 1972).

[16] J. A. van Casteren, A problem of Sz.-Nagy, *Acta Sci. Math.*, **42** (1980), 189—194.

DEPARTMENT OF MATHEMATICS
TECHNION — ISRAEL INSTITUTE OF TECHNOLOGY
32000 HAIFA, ISRAEL

Current address:
SCHOOL OF MATH. SCIENCES
TEL-AVIV UNIVERSITY
RAMAT AVIV
69978 TEL-AVIV, ISRAEL

Acta Sci. Math., 50 (1986), 213—223

The distance between unitary orbits of normal operators

KENNETH R. DAVIDSON[1])

The problem of computing the distance between unitary orbits of operators is an important and difficult problem, even in the finite dimensional case. These problems have a long history, and we will mention some of the important results for the operator norm only.

In 1912, WEYL [21] proved that given two Hermitian matrices A and B with spectrum $a_1 \leqq a_2 \leqq \ldots \leqq a_n$ and $b_1 \leqq b_2 \leqq \ldots \leqq b_n$ (repeated according to multiplicity) respectively, then

$$\operatorname{dist}(\mathcal{U}(A), \mathcal{U}(B)) = \max_{1 \leqq i \leqq n} |a_i - b_i|.$$

This distance is clearly attained by a commuting pair of diagonal matrices in $\mathcal{U}(A)$ and $\mathcal{U}(B)$, respectively.

The normal case has received much attention, but the final answer is still not known. However, the natural analogue for the right hand side is obtained by looking at *commuting* pairs. If A and B commute, they are simultaneously diagonalizable which results in a pairing of eigenvalues $\{a_i, b_i\}$, $1 \leqq i \leqq n$, and $\|A-B\| = \max |a_i - b_i|$. This is minimized if the pairing is optimal. This suggests the *spectral distance*

$$\delta(A, B) = \min_{\pi} \max_{1 \leqq i \leqq n} |a_i - b_{\pi(i)}|$$

where π runs over all permutations. Recently, it has been shown [5] that there is a universal constant c independent of dimension such that

$$\delta(A, B) \geqq \operatorname{dist}(\mathcal{U}(A), \mathcal{U}(B)) \geqq c^{-1} \delta(A, B).$$

A number of cases in which equality (i.e. $c=1$) exists are known: unitaries [4], self-adjoint and skew-adjoint [18], and scalar multiples of unitaries [6].

In infinite dimensions, unitary orbits have received much attention. In this case, $\mathcal{U}(A)$ is rarely closed. However, for normals, there is a nice description of

Received February 27, 1984.

[1]) Research partially supported by Natural Sciences and Engineering Research Council of Canada (NSERC).

$\overline{\mathscr{U}(A)}$ in terms of a crude multiplicity function [10]. This, in turn, can be interpreted in terms of the spectrum of A modulo various (closed 2-sided) ideals of $\mathscr{B}(H)$ [8]. For more general operators, the invariants for $\overline{\mathscr{U}(A)}$ can be very complicated, and on this there is a large literature ([16], [9], [19], [11]). An important problem in operator theory and C^* algebras has been the relation between the unitary orbit and ideal perturbations. The grandfather result is the Weyl—von Neumann Theorem and it has many important successors [3], [8], [19], [20], [12].

There is less information known about distances between unitary orbits. Some of this information has been obtained in an effort to give quantitative estimates in the study of compact perturbations following [7]. For example, Berg [3] gave concrete estimates for the distance between unitary orbits of direct sums of normals and weighted shifts. Later, Herrero [15] gave very good estimates of the distance between unitary orbits of power partial isometries by improving on Berg's technique. Finally, the problem for pairs of self-adjoint operators has recently been solved [1]. They define a spectral distance in terms of the crude multiplicity function, and show that this is exactly the distance between the unitary orbits. Furthermore, this distance is achieved by commuting diagonal operators in the closure of the orbits.

This is the starting point for the work of this paper. We show that the same spectral distance is the right one for normal operators. This distance is the infimum of $\|A-B\|$ for commuting pairs in the closed orbits, and this distance is attained by a pair of commuting diagonal operators. When A has no isolated eigenvalues of finite multiplicity, this is exactly the distance between orbits. In general, we obtain that

$$\operatorname{dist}(\mathscr{U}(A), \mathscr{U}(B)) \geqq c^{-1}\delta(A, B)$$

where c is the same constant as in [5]. The general problem of determining if $c=1$ reduces to the separable case with finite spectra. We do not know if a positive answer in the finite dimensional case would imply the same in the separable case. However, we believe that any proof would almost surely generalize.

1. Preliminaries

Let $\mathfrak{H}$ be a Hilbert space, and let $\mathscr{B}(\mathfrak{H})$ denote the algebra of bounded linear operators on $\mathfrak{H}$. Given an operator A, $\mathscr{U}(A)$ denotes the set $\{UAU^*\colon U \text{ unitary in } \mathscr{B}(\mathfrak{H})\}$ and $\overline{\mathscr{U}(A)}$ denotes its closure. For $\mathfrak{M}$ a closed subspace of $\mathfrak{H}$, $\dim \mathfrak{M}$ is the cardinality of an orthonormal basis for $\mathfrak{M}$. Let h denote the dimension of $\mathfrak{H}$. For each infinite cardinal $\alpha \leqq h$, let $\mathscr{I}_\alpha$ denote the closed two sided ideal generated by $\{T \in \mathscr{B}(\mathfrak{H})\colon \dim \overline{\operatorname{Ran} T} < \alpha\}$. Let $\sigma_\alpha(A)$ denote $\sigma(A+\mathscr{I}_\alpha)$ as an element of the quotient C^* algebra $\mathscr{B}(\mathfrak{H})/\mathscr{I}_\alpha$. In particular, $\mathscr{I}_{\aleph_0}$ is the set of compact operators and $\sigma_{\aleph_0}(A)=\sigma_e(A)$, the essential spectrum. Also, let $\sigma_0(A)$ denote the isolated points of finite multiplicity in $\sigma(A)$, known as the normal eigenvalues of A. For convenience of notation, we write $\sigma_1(A)$ for $\sigma(A)$.

Let A be a normal operator, and let $E_A(\cdot)$ denote its spectral measure. For $r>0$, λ in $\mathbf{C}$, the disc of radius r about λ is denoted $D_r(\lambda)$. In [10], a *crude multiplicity function* is defined for normal operators by

$$\alpha(\lambda)=\inf\{\operatorname{rank} E_A(D_r(\lambda)):\ r>0\}.$$

It is shown there that two normal operators on a separable space have the same closed unitary orbit if and only if they have the same crude multiplicity function. This is easily generalized to Hilbert spaces of arbitrary cardinality, and is a special case of HADWIN's Theorem 3.14 [12]. It is not difficult to see (cf. [8]) that if α is an infinite cardinal, then $\{\lambda:\ \alpha(\lambda)\geqq\alpha\}$ equals $\sigma_\alpha(A)$. For $\alpha(\lambda)=n$ to be finite, non-zero, λ must be an isolated eigenvalue of multiplicity n. Thus the theorem of Gellar—Page and Hadwin can be formulated:

Proposition 1.1. *Two normal operators A and B have the same closed unitary orbits if and only if $\sigma_\alpha(A)=\sigma_\alpha(B)$ for each infinite cardinal, and $\sigma_0(A)=\sigma_0(B)$ including multiplicity.*

We are now ready to discuss the spectral distance formula of AZOFF and DAVIS [1]. In the finite case, one might also define a spectral distance

$$\varrho(A,B)=\sup_{F\text{ finite}}\inf\{r:\ \operatorname{rank} E_A(F)\leqq\operatorname{rank} E_B(F_r)\ \text{ and }\ \operatorname{rank} E_B(F)\leqq\operatorname{rank} E_A(F_r)\}$$

where $F_r=\{\lambda:\ \operatorname{dist}(\lambda,F)\leqq r\}$. A simple application of the Marrige Lemma [14] shows that indeed $\varrho=\delta$. So we define our spectral distance as follows:

Definition 1.2. $\delta(A,B)$ is the infimum of real numbers $r>0$ such that

$$\operatorname{rank} E_A(F)\leqq\operatorname{rank} E_B(F_r)\quad\text{and}\quad\operatorname{rank} E_B(F)\leqq\operatorname{rank} E_A(F_r)$$

for every compact subset F of $\mathbf{C}$.

We wish to relate this formula to the various spectra. Let $d_H(X,Y)$ denote the Hausdorff distance

$$d_H(X,Y)=\max\{\sup_{x\in X}\operatorname{dist}(x,Y),\ \sup_{y\in Y}\operatorname{dist}(y,X)\}.$$

Let $\operatorname{Fin}(X)$ denote the collection of finite subsets of X. Let $\delta_f(A,B)$ denote the maximum of

$$\sup_{F\in\operatorname{Fin}\sigma_0(A)}\inf\{r:\ \operatorname{rank} E_A(F)\leqq\operatorname{rank} E_B(F_r)\}$$

and the corresponding term with A and B interchanged. (Should $\sigma_0(A)$ be empty, this term is defined to be zero.) We obtain:

Proposition 1.3. *Let A and B be normal operators. Then*

$$\delta(A,B) = \max\{d_f(A,B), \sup_{\aleph_0 \leqq \alpha \leqq h} d_H(\sigma_\alpha(A), \sigma_\alpha(B))\}.$$

Proof. Let α be an infinite cardinal, and suppose λ belongs to $\sigma_\alpha(A)$. Then $\operatorname{rank} E_A(D_\varepsilon(\lambda)) \geqq \alpha$ for all $\varepsilon > 0$. Thus if $r = \delta(A,B)$, we have $\operatorname{rank} E_B(D_{r+\varepsilon}(\lambda)) \geqq \alpha$ for all $\varepsilon > 0$. Hence $\sigma_\alpha(B)$ intersects $\overline{D_r(\lambda)}$, so $\operatorname{dist}(\lambda, \sigma_\alpha(B)) \leqq r$. Letting α and λ run over all possibilities, and then interchanging the role of A and B, we obtain

$$\delta(A,B) \geqq \sup_{\aleph_0 \leqq \alpha \leqq h} d_H(\sigma_\alpha(A), \sigma_\alpha(B)).$$

By definition, $\delta(A,B) \geqq \delta_f(A,B)$, so $\delta(A,B)$ is greater than the right hand side, say s.

Conversely, let F be a compact subset of $\mathbf{C}$, and let $\alpha = \operatorname{rank} E_A(F)$. If α is infinite, then $\sigma_\alpha(A)$ intersects F. Thus $\sigma_\alpha(B)$ intersects F_s, and thus $\operatorname{rank} E_B(F_{s+\varepsilon}) \geqq \alpha$ for all $\varepsilon > 0$. If α is finite and F is contained in $\sigma_0(A)$, the definition of $\delta_f(A, B)$ gives $\alpha \leqq \operatorname{rank} E_B(F_{s+\varepsilon})$ for $\varepsilon > 0$. Finally, if α is finite but F is not contained in $\sigma_0(A)$, then F intersects $\sigma_e(A)$. So as in the infinite case, $\alpha \leqq \aleph_0 \leqq \operatorname{rank} E_B(F_{s+\varepsilon})$ for all $\varepsilon > 0$. It follows that $\delta(A,B) \leqq s$, and equality is obtained.

Remark 1.4. If $\sigma(A) = \sigma_e(A)$, then

$$\delta(A,B) = \sup_{1 \leqq \alpha \leqq h} d_H(\sigma_\alpha(A), \sigma_\alpha(B)).$$

Here α runs over the infinite cardinals and $\alpha = 1$, where $\sigma_1(B) = \sigma(B)$ by definition. To see this, note that if F is a finite subset of $\sigma_0(B)$, then $\operatorname{rank} E_A(F_r)$ equals zero or is infinite. So $\inf\{r\colon \operatorname{rank} E_A(F_r) \geqq \operatorname{rank} E_B(F)\} = \inf\{r\colon F_r \cap \sigma(A) \neq \emptyset\}$. Hence $\delta_f(A,B) = \sup_{\lambda \in \sigma_0(B)} \operatorname{dist}(\lambda, \sigma(A))$.

2. Lower bounds

Proposition 2.1. *If A and B are normal operators on $\mathfrak{H}$, then*

$$\operatorname{dist}(\mathcal{U}(A), \mathcal{U}(B)) \geqq \sup_{1 \leqq \alpha \leqq h} d_H(\sigma_\alpha(A), \sigma_\alpha(B)).$$

Proof. Let $\alpha = 1$ or some infinite cardinal, and let λ belong to $\sigma_\alpha(A)$. Then $\operatorname{rank} E_A(D_\varepsilon(\lambda)) \geqq \alpha$ for all $\varepsilon > 0$. If B' belongs to $\mathcal{U}(B)$ and $\|A - B'\| = s$, we will show that $\operatorname{rank} E_B(D_{s+\varepsilon}(\lambda)) \geqq \alpha$ for all $\varepsilon > 0$. Otherwise for some $\varepsilon > 0$, there is a unit vector x in $\operatorname{Ran} E_A(D_{\varepsilon/2}(\lambda)) \cap \operatorname{Ran} E_{B'}(D_{s+\varepsilon}(\lambda))^\perp$. This gives

$$\|A - B'\| \geqq \|(A - B')x\| \geqq \|(B' - \lambda)x\| - \|(A - \lambda)x\| \geqq s + \varepsilon - \varepsilon/2 = s + \varepsilon/2.$$

Hence $\sigma_\alpha(B)$ intersects $\overline{D_s(\lambda)}$, and thus $\operatorname{dist}(\lambda, \sigma_\alpha(B)) \leqq s$. By symmetry, $d_H(\sigma_\alpha(A), \sigma_\alpha(B)) \leqq \operatorname{dist}(\mathcal{U}(A), \mathcal{U}(B))$ for all α.

Corollary 2.2. *If A and B are normal, and $\sigma(A)=\sigma_e(A)$, then*

$$\operatorname{dist}(\mathcal{U}(A), \mathcal{U}(B)) \geqq \delta(A, B).$$

Proposition 2.3. *If A and B are commuting normal operators, then*

$$\|A-B\| \geqq \delta(A, B).$$

Proof. By Propositions 1.3 and 2.1, we need only show that $\|A-B\| \geqq \delta_f(A, B)$. Suppose F is a finite subset of $\sigma_0(A)$, $s>0$, and $\operatorname{rank} E_A(F) > \operatorname{rank} E_B(F_s)$. It suffices to show that $\|A-B\| \geqq s$. Now $E_A(F)$ belongs to $W^*(A)=\{A\}''$, and thus commutes with B. So the restrictions A_0 and B_0 of A and B to $E_A(F)\mathfrak{H}$ are commuting normal operators. Hence A_0 and B_0 are simultaneously diagonalizable. The spectrum of A_0 lies in F, but B_0 has at least one eigenvalue outside F_s, which is paired with some eigenvalue of A_0. Thus

$$\|A-B\| \geqq \|A_0-B_0\| \geqq s.$$

The main result of this section is an easy corollary of the result of [5].

Theorem 2.4. *There is a universal constant $c>0$ so that for every pair of normal operators A and B acting on the same space,*

$$\operatorname{dist}(\mathcal{U}(A), \mathcal{U}(B)) \geqq c^{-1}\delta(A, B).$$

Proof. By Propositions 1.3 and 2.1, it suffices to show that $\|A-B\| \geqq c^{-1}\delta_f(A, B)$. So let F be a finite subset of $\sigma_0(A)$, $s>0$, and $\operatorname{rank} E_A(F) > \operatorname{rank} E_B(F_s)$. Let $\mathfrak{K}=E_A(F)$, $\mathfrak{L}=E_B(F_s)^\perp\mathfrak{H}$, $\tilde{A}=A|\mathfrak{K}$ and $\tilde{B}=B|\mathfrak{L}$. Let $C=E_A(F)E_B(F_s)^\perp$ be thought of as an operator from $\mathfrak{L}$ to $\mathfrak{K}$. Then

$$\|\tilde{A}Q-Q\tilde{B}\| = \|E_A(F)(A-B)E_B(F_s)^\perp\| \leqq \|A-B\|.$$

By Theorem 4.2 of [5], we get

$$\|A-B\| \geqq sc^{-1}\|Q\|.$$

However, since $\operatorname{codim} \mathfrak{L} = \operatorname{rank} E_B(F_s) < \dim \mathfrak{K}$, $\mathfrak{L}$ and $\mathfrak{K}$ intersect, and thus $\|Q\|=1$. This completes the proof.

3. Best commuting approximants

Now we construct closest possible diagonal operators in the unitary orbits. There is a technical matching problem that has to be solved, and this will be left to the next section.

In the following proof, we make use of this fact: If X is a compact subset of the plane, there is a Borel function f mapping $\mathbf{C}$ to X so that $|z-f(z)|=d(z)\equiv \operatorname{dist}(z, X)$. To obtain such an f, let $F(z)=\{x\in X\colon |z-x|=d(z)\}$. Let

$$\theta(z) = \min \{\operatorname{Arg}(x-z)\colon x\in F(z)\}$$

where $\operatorname{Arg}(w)$ belongs to $[0, 2\pi)$ such that $w=|w|\exp(i\operatorname{Arg} w)$. Let $f(z)=$ $=z+d(z)\exp(i\theta(z))$. It is readily verified that f is Borel as required.

Theorem 3.1. *Suppose A and B are normal operators on $\mathfrak{H}$. There are commuting diagonal operators A' and B' in $\overline{\mathcal{U}(A)}$ and $\overline{\mathcal{U}(B)}$ respectively such that $\|A'-B'\|=$ $=\delta(A, B)$.*

Proof. Let $r=\delta(A, B)$. For each infinite cardinal $\alpha\leqq h$, let A_α be a diagonal normal operator on a Hilbert space $\mathfrak{H}_\alpha$ of cardinality α with $\sigma(A_\alpha)=\sigma_\alpha(A_\alpha)=\sigma_\alpha(A)$. Similarly, let B_α be a diagonal normal operator on $\mathfrak{H}_\alpha$ with $\sigma(B_\alpha)=\sigma_\alpha(B_\alpha)=\sigma_\alpha(B)$. Let f_α be a Borel map of $\sigma_\alpha(A)$ into $\sigma_\alpha(B)$ such that $|f_\alpha(a)-a|\leqq r$ for all a in $\sigma_\alpha(A)$. Similarly, let g_α be a Borel map of $\sigma_\alpha(B)$ into $\sigma_\alpha(A)$ such that $|g_\alpha(b)-b|\leqq r$. These maps exist since $d_H(\sigma_\alpha(A), \sigma_\alpha(B))\leqq r$. Consider the diagonal normal operators $\hat{A}_\alpha=A_\alpha\oplus g_\alpha(B_\alpha)$ and $\hat{B}_\alpha=f_\alpha(A_\alpha)\oplus B_\alpha$ acting on $\mathfrak{H}_\alpha\oplus\mathfrak{H}_\alpha$. This construction guarantees that $\dim(\mathfrak{H}_\alpha\oplus\mathfrak{H}_\alpha)=\alpha$, $\sigma(\hat{A}_\alpha)=\sigma_\alpha(\hat{A}_\alpha)=\sigma_\alpha(A)$, $\sigma(\hat{B}_\alpha)=\sigma_\alpha(\hat{B}_\alpha)=\sigma_\alpha(B)$, and $\|\hat{A}_\alpha-\hat{B}_\alpha\|\leqq r$.

Consider $\mathfrak{H}=\sum_{\aleph_0\leqq\alpha\leqq h}(\mathfrak{H}_\alpha\oplus\mathfrak{H}_\alpha)$, and the diagonal normal operators $\hat{A}=\oplus\sum_{\aleph_0\leqq\alpha\leqq h}\hat{A}_\alpha$ and $\hat{B}=\oplus\sum_{\aleph_0\leqq\alpha\leqq h}\hat{B}_\alpha$. These are commuting diagonal operators such that $\|\hat{A}-\hat{B}\|\leqq r$, $\sigma_\alpha(\hat{A})=\sigma_\alpha(A)$ and $\sigma_\alpha(\hat{B})=\sigma_\alpha(B)$ for all infinite cardinals, and $\sigma_0(\hat{A})=\emptyset=\sigma_0(\hat{B})$.

To obtain the desired operators, we need to add on summands to give the isolated eigenvalues of finite multiplicity. We will produce commuting diagonal normal operators A_f and B_f on a separable space such that $\|A_f-B_f\|\leqq r$, $\sigma_e(A_f)\subseteq$ $\subseteq\sigma_e(A)$ (and $\sigma_e(B_f)\subseteq\sigma_e(B)$) and $\sigma_0(A_f)\setminus\sigma_e(A)$ and $\sigma_0(A)$ (correspondingly $\sigma_0(B_f)\setminus\sigma_e(B)$ and $\sigma_0(B)$) coincide including multiplicity. Once this is accomplished, let $A'=\hat{A}\oplus A_f$ and $B'=\hat{B}\oplus B_f$. These are commuting diagonal normals such that $\|A'-B'\|\leqq r$. Furthermore, $\sigma_\alpha(A')=\sigma_\alpha(A)$ for infinite cardinals and $\sigma_0(A')=$ $=\sigma_0(A)$ including multiplicity, so by Proposition 1.1, A' belongs to $\overline{\mathcal{U}(A)}$. Similarly, B' belongs to $\overline{\mathcal{U}(B)}$.

To produce A_f and B_f, we must find a matching of the points in $\sigma_0(A)$ to points in $\sigma(B)$ within distance r. Points close to $\sigma_e(B)$ can be "absorbed", but points further from $\sigma_e(B)$ must match up including multiplicity with points in $\sigma_0(B)$; and vice versa. To do this, we need a curious Marriage Lemma type theorem. This theorem will be stated and proved in the next section. First we obtain the appropriate setting for our problem, and finish the proof modulo this combinatorical theorem.

Let $\mathscr{A}_0$ denote the points of $\sigma_0(A)$ repeated according to multiplicity which are distance greater than r from $\sigma_e(B)$. Similarly let $\mathscr{B}_0$ be the corresponding subset of $\sigma_0(B)$. Define a relation R on $\sigma_0(A)\times\sigma_0(B)$ by aRb if and only if $|a-b|\leqq r$. For each $\mathscr{A}$ in $\mathrm{Fin}(\mathscr{A}_0)$, the set $F(\mathscr{A})=\{b\in\sigma_0(B)\colon aRb$ for some a in $\mathscr{A}\}$ is finite. And since $\delta(A,B)\geqq\delta_f(A,B)$, we have $|\mathscr{A}|\leqq|F(\mathscr{A})|<\infty$. Similarly, set $G(\mathscr{B})=\{a\in\sigma_0(A)\colon aRb$ for some b in $\mathscr{B}\}$ for each $\mathscr{B}$ in $\mathrm{Fin}(\mathscr{B}_0)$. Again, $|\mathscr{B}|\leqq|G(\mathscr{B})|<\infty$.

This relation satisfies the hypotheses of the Combinatorial Lemma 4.1. Thus there are sets $\mathscr{A}_1$ and $\mathscr{B}_1$ such that $\mathscr{A}_0\subseteqq\mathscr{A}_1\subseteqq\sigma_0(A)$ and $\mathscr{B}_0\subseteqq\mathscr{B}_1\subseteqq\sigma_0(B)$, and a *bijection* $f\colon \mathscr{A}_1\to\mathscr{B}_1$ such that $|a-f(a)|\leqq r$ for every a in $\mathscr{A}_1$. Let $\mathscr{A}_2=\sigma_0(A)\setminus\mathscr{A}_1$ and choose a function $f_2\colon \mathscr{A}_2\to\sigma_e(B)$ such that $|f_2(a)-a|\leqq r$ for all a in $\mathscr{A}_2$. Similarly, define $\mathscr{B}_2$ and $g_2\colon \mathscr{B}_2\to\sigma_e(A)$.

Define diagonal normal operators as follows:

$$\begin{aligned}
A_1 &= \mathrm{diag}\,\{a_n\colon a_n\in\mathscr{A}_1\}, & B_1 &= \mathrm{diag}\,\{f(a_n)\colon a_n\in\mathscr{A}_1\},\\
A_2 &= \mathrm{diag}\,\{a_n\colon a_n\in\mathscr{A}_2\}, & B_\infty &= \mathrm{diag}\,\{f_2(a_n)\colon a_n\in\mathscr{A}_2\},\\
A_\infty &= \mathrm{diag}\,\{g_2(b_n)\colon b_n\in\mathscr{B}_2\}, & B_2 &= \mathrm{diag}\,\{b_n\colon b_n\in\mathscr{B}_2\},\\
A_f &= A_1\oplus A_2\oplus A_\infty, & B_f &= B_1\oplus B_\infty\oplus B_2.
\end{aligned}$$

The properties of f, f_2 and g_2 show that $\|A_f-B_f\|\leqq r$, $\sigma_0(A_1\oplus A_2)$ agrees with $\sigma_0(A)$ including multiplicity, and $\sigma_e(A_f)\cup\sigma_0(A_\infty)\subseteqq\sigma_e(A)$. The corresponding statements for B_f hold also. So A_f and B_f are commuting diagonal operators with the required properties. This completes the proof.

Corollary 3.2. *If A and B are normal operators, then*

$$\mathrm{dist}\,(\mathscr{U}(A),\mathscr{U}(B))\leqq\delta(A,B).$$

Corollary 3.3. *If A and B are normal operators, and $\sigma(A)=\sigma_e(A)$, then*

$$\mathrm{dist}\,(\mathscr{U}(A),\mathscr{U}(B))=\delta(A,B).$$

Remark 3.4. The equality $\mathrm{dist}\,(\mathscr{U}(A),\mathscr{U}(B))=\delta(A,B)$ is readily verified for several classes of normals: the self adjoint case [1] can also be proven using the technique of [6], as can the case of scalar multiples of unitaries. The technique of [5] works for a self-adjoint A and skew-adjoint B. Just remember that one needs only worry about $\delta_f(A,B)$.

4. The combinatorial lemma

The result that we need is an infinite analogue of the Marriage Lemma [14]. There are a number of infinite versions of this theorem, notably [13] and [17], and our proof is very similar. However, the set up we require seems sufficiently peculiar that no known theorem applies directly.

Let $\mathscr{A}$ and $\mathscr{B}$ be sets, and let R be a relation on $\mathscr{A}\times\mathscr{B}$. For A a subset of $\mathscr{A}$, define $F(A)=\{b\in\mathscr{B}: aRb$ for some a in $A\}$. Similarly, for B a subset of $\mathscr{B}$, define $G(B)=\{a\in\mathscr{A}: aRb$ for some b in $B\}$.

Lemma 4.1. *Let $\mathscr{A}$ and $\mathscr{B}$ be sets with distinguished subsets $\mathscr{A}_0$ and $\mathscr{B}_0$; and let R be a relation on $\mathscr{A}\times\mathscr{B}$ with F and G defined as above. Suppose that*

1) $|A|\leqq|F(A)|<\infty$ *for every A in* $\mathrm{Fin}(\mathscr{A}_0)$,

1') $|B|\leqq|G(B)|<\infty$ *for every B in* $\mathrm{Fin}(\mathscr{B}_0)$.

Then there are sets $\mathscr{A}_1$ and $\mathscr{B}_1$ such that $\mathscr{A}_0\subseteqq\mathscr{A}_1\subseteqq\mathscr{A}$ and $\mathscr{B}_0\subseteqq\mathscr{B}_1\subseteqq\mathscr{B}$, and a bijection $f\colon\mathscr{A}_1\to\mathscr{B}_1$ such that $aRf(a)$ for every a in $\mathscr{A}_1$.

Proof. Call a non-empty set A in $\mathrm{Fin}(\mathscr{A}_0)$ *strict* if $|F(A)|=|A|$. The restriction of R to $A\times F(A)$ satisfies 1). So the Marriage Lemma [14] gives a bijection $f\colon A\to F(A)$ such that $aRf(a)$ for every a in A.

Note also that when A is strict, $\mathscr{A}\setminus A$ and $\mathscr{B}\setminus F(A)$ with distinguished sets $\mathscr{A}_0\setminus A$ and $\mathscr{B}_0\setminus F(A)$ still satisfy 1) and 1'). For if A' is a finite subset of $\mathscr{A}_0\setminus A$, then

$$|F(A')\setminus F(A)| = |F(A'\cup A)\setminus F(A)| \geqq |A'\cup A|-|F(A)| = |A'|.$$

Also if B' is a finite subset of $\mathscr{B}_0\setminus F(A)$, then $G(B)$ is disjoint from A, so $|B'|\leqq\leqq|G(B')|=|G(B')\setminus A|$.

Similarly, if B is a strict subset of $\mathscr{B}_0$, there is likewise a bijection $f\colon G(B)\to B$ such that $f(a)Ra$ for all a in $G(B)$. And $\mathscr{A}\setminus G(B)$ and $\mathscr{B}\setminus B$ satisfy 1) and 1').

Use the Axiom of Choice to well order $\mathscr{A}_0\cup\mathscr{B}_0$. Starting out of (A, B) we define (A_α, B_α) by transfinite induction. At stage α, we have a collection $\{(\mathscr{A}_\beta; \mathscr{B}_\beta)\colon \beta<\alpha\}$ of pairs satisfying 1), 1') and $\mathscr{A}_\beta\supset\mathscr{A}_{\beta'}$ and $\mathscr{B}_\beta\supset\mathscr{B}_{\beta'}$ if $\beta<\beta'<\alpha$. When α is a limit ordinal, set $\mathscr{A}_\alpha=\bigcap_{\beta<\alpha}\mathscr{A}_\beta$ and $\mathscr{B}_\alpha=\bigcap_{\beta<\alpha}\mathscr{B}_\beta$. Since 1) and 1') deal with finite sets, it is easy to verify that the hold for the intersection.

If $\alpha=\beta+1$ is a successor ordinal, then

(a) If there are strict subsets of $\mathscr{A}_0\cap\mathscr{A}_\beta$, choose a strict A and obtain a pair $(A, F(A))$.

(b) If there are no strict subsets of $\mathscr{A}_0\cap\mathscr{A}_\beta$, but there are strict subsets of $\mathscr{B}_0\cap\mathscr{B}_\beta$, choose a strict B and form a pair $(G(B); B)$.

(c) If there are no strict subsets, let a (or b) be the least element of the well ordering of $\mathscr{A}_0\cup\mathscr{B}_0$ which belong to $\mathscr{A}_\beta\cup\mathscr{B}_\beta$. Take any b (or a) such that aRb, and form the pair $(\{a\}; \{b\})$.

In each case we obtain a pair (A, B) with $|A|=|B|<\infty$, and, following earlier remarks, a bijection $f\colon A\to B$ such that $aRf(a)$ for a in A. Furthermore, we set $\mathscr{A}_\alpha=\mathscr{A}_\beta\setminus A$ and $\mathscr{B}_\alpha=\mathscr{B}_\beta\setminus B$. By the remarks at the beginning of the proof, $\mathscr{A}_\alpha$ and $\mathscr{B}_\alpha$ satisfy 1) and 1') in cases (a) and (b). They also hold in case (c), for if A is a finite subset of $\mathscr{A}_\alpha\cap\mathscr{A}_0$, then since A is not strict,

$$|F(A)\setminus\{b\}| \geqq |F(A)|-1 \geqq |A|.$$

The same holds for finite subsets of $\mathscr{B}_\alpha\cap\mathscr{B}_0$.

This procedure terminates at some ordinal α with $|\alpha|\leqq|\mathscr{A}_0\cup\mathscr{B}_0|$. The result is a disjoint collection of finite pairs (A_i, B_i) which exhaust $\mathscr{A}_0$ and $\mathscr{B}_0$. On these sets, functions $f=f_i$ have been constructed so that $aRf(a)$ for all a in A_i. Set $\mathscr{A}_1=\bigcup A_i$ and $\mathscr{B}_1=\bigcup B_i$. The union f of the f_i's is the required bijection.

5. Further remarks

The results of section 2 show that the problem of determining if the constant c equals 1 is basically a finite dimensional problem, for the difficulty lies solely in the $\delta_f(A, B)$ term. A quantitative way of phrasing the key issue is

Question. If A and B are normal operators (on a separable space), F is a finite subset of $\sigma_0(A)$, $s>0$, and $\operatorname{rank} E_B(F_s)<\operatorname{rank} E_A(F)$, then is $\|A-B\|\geqq s$?

The reason one can reduce to the separable case is the following. Suppose A and B are normal with $\|A-B\|<\delta(A, B)$. Let $\mathfrak{M}$ be a separable reducing subspace of A containing the spectral subspace $E_A\big(\sigma(A)\setminus\sigma_{\aleph_1}(A)\big)$ such that $\sigma_e(A|\mathfrak{M})=\sigma_e(A)$, and let $\mathfrak{N}$ be a corresponding subspace for B. Let $\mathfrak{K}$ be the smallest reducing subspace for $C^*(A, B)$ containing both $\mathfrak{M}$ and $\mathfrak{N}$. This subspace is separable. Let $A'=A|\mathfrak{K}$, $A''=A|\mathfrak{K}^\perp$, $B'=B|\mathfrak{K}$ and $B''=B|\mathfrak{K}^\perp$. Then $\sigma_0(A')=\sigma_0(A)$ and $\sigma_e(A')=\sigma_e(A)$; and $\sigma(A'')=\sigma_{\aleph_1}(A'')=\sigma_{\aleph_1}(A)$. Similar relations hold for B. Furthermore,

$$\delta(A'', B'') = \sup_{\aleph_1\leqq\alpha\leqq h} d_H\big(\sigma_\alpha(A''), \sigma_\alpha(B'')\big) = \sup_{\aleph_1\leqq\alpha\leqq h} d_H\big(\sigma_\alpha(A), \sigma_\alpha(B)\big) \leqq$$

$$\leqq \|A''-B''\| \leqq \|A-B\|$$

and

$$\delta(A', B') = \max\big\{\delta_f(A, B),\ d_H\big(\sigma_e(A), \sigma_e(B)\big)\big\}.$$

Thus $\delta(A, B)=\max\{\delta(A', B'), \delta(A'', B'')\}$ and $\|A-B\|=\max\{\|A'-B'\|, \|A''-B''\|\}$. So it must be the case that

$$\delta(A', B') = \delta(A, B) > \|A-B\| \geqq \|A'-B'\|.$$

This reasonsing also leads to the conclusion that any constant c valid in Theorem 2.4 for separable spaces is valid in general.

Next, it is easy to approximate A and B arbitrarily well by normals of finite spectrum. So one may assume that $\sigma(A)$ and $\sigma(B)$ are finite. This looks almost finite dimensional now.

Question. Can one show that any constant c valid in Theorem 2.4 for all finite rank normals works in general?

I am confident that any proof valid in the finite case will extend to the separable one, but knowing this in advance would be nice.

Finally, a special case subsuming much of what is known and is perhaps easier than the general case is the situation $A=A^*$ and arbitrary normal B. Does this case have $c=1$?

Bibliography

[1] E. Azoff and C. Davis, On distances between unitary orbits of self-adjoint operators, *Acta Sci. Math.*, **47** (1984), 419—439.

[2] I. D. Berg, An extension of the Weyl—von Neumann Theorem to normal operators, *Trans. Amer. Math. Soc.*, **160** (1971), 365—371.

[3] I. D. Berg, Index theory for perturbations of direct sums of normal operators and weighted shifts, *Canad. J. Math.*, **30** (1978), 1152—1165.

[4] R. Bhatia and C. Davis, A bound for the spectral variation of a unitary operator, *Linear and Multilinear Algebra*, **15** (1984), 71—76.

[5] R. Bhatia, C. Davis and A. McIntosh, Perturbations of spectral subspaces and solution of linear operator equations, *Linear Algebra Appl.*, **52/53** (1983), 45—67.

[6] R. Bhatia and J. Holbrook, Short normal paths and spectral variation, *Proc. Amer. Math. Soc.*, **94** (1985), 377—382.

[7] L. Brown, R. Douglas and P. Fillmore, Unitary equivalence modulo the compact operators, and extensions of C^* algebras, in: *Proceedings of a Conference on Operator Theory* (Halifax, 1973), Lecture Notes in Math. vol. 345, Springer-Verlag (Berlin, 1973); pp. 58—123.

[8] G. Edgar, J. Ernest and S. G. Lee, Weighing operator spectra, *Indiana Univ. Math. J.*, **21** (1971), 61—80.

[9] J. Ernest, Charting the operator terrain, *Mem. Amer. Math. Soc.*, **171** (1976).

[10] R. Gellar and L. Page, Limits of unitarily equivalent normal operators, *Duke Math. J.*, **41** (1974), 319—322.

[11] D. Hadwin, An operator valued spectrum, *Indiana Univ. Math. J.*, **26** (1977), 329—340.

[12] D. Hadwin, Non-separable approximate equivalence, *Trans. Amer. Math. Soc.*, **266** (1981), 203—231.

[13] M. Hall, Jr., Distinct representatives of subsets, *Bull. Amer. Math. Soc.*, **54** (1948), 922—928.

[14] P. Hall, On representation of subsets, *J. London Math. Soc.*, **10** (1935), 26—30.

[15] D. Herrero, Unitary orbits of power partial isometries and approximation by block diagonal nilpotents, in: *Topics in Modern Operator Theory*, Birkhäuser-Verlag (Basel, 1981); pp. 171—210.

[16] C. PEARCY and N. SALINAS, The reducing essential matricial spectra of an operator, *Duke Math. J.*, **42** (1975), 423—434.

[17] R. RADO, The transfinite case of Hall's theorem on representatives, *J. London Math. Soc.*, **42** (1967), 321—324.

[18] V. S. SUNDER, Distance between normal operators, *Proc. Amer. Math. Soc.*, **84** (1982), 483—484.

[19] D. VOICULESCU, A non-commutative Weyl—von Neumann theorem, *Rev. Roumaine Math. Pures Appl.*, **21** (1976), 97—113.

[20] D. VOICULESCU, Some results on norm ideal perturbations of Hilbert space operators, *J. Oper. Theory*, **2** (1979), 3—37.

[21] H. WEYL, Der asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen, *Math. Ann.*, **71** (1912), 441—479.

DEPARTMENT OF PURE MATHEMATICS
UNIVERSITY OF WATERLOO
WATERLOO, ONTARIO, CANADA N2L 3G1

Acta Sci. Math., 50 (1986), 225—230

A note on integral operators

BEN DE PAGTER

In 1974 A. V. BUHVALOV [1] proved that the set of all absolute integral operators from an ideal L into an ideal M of measurable functions is equal to the band generated by the integral operators of finite rank. A detailed discussion of this theorem can be found in [3] (Chapter 13). This result was proved under the additional hypothesis that the normal integrals on the domain L (i.e., linear functionals φ that can be written as $\varphi=\varphi_1-\varphi_2$, where φ_1 and φ_2 are positive order continuous linear functionals on L) separate the points, which guarantees the existence of sufficiently many non-zero integral operators. Some time ago it was conjectured by A. C. Zaanen that Buhvalov's result remains valid without the assumption concerning the normal integrals. In the present paper we prove that the conjecture is true. In particular it will be shown that if L is an ideal of measurable functions not possessing any non-zero normal integrals, then L cannot be the domain of any non-zero integral operator. This last situation occurs for example if (Y, Σ, ν) is a finite measure space not containing any atoms and if we take for L any of the ideals $L_p(Y, \nu)$ $(0<p<1)$, $L_{(1,\infty)}(Y, \nu)$ (the weak L_1 functions) or the space $L_0(Y, \nu)$ of all ν-measurable functions on X.

In this paper we restrict ourselves to considering real measurable functions only; since all the results can be extended easily to the complex case by means of complexification (see e.g. [3], sections 91 and 92).

1. We start with some notation and terminology. We refer to the book [2] for any unexplained terminology concerning Riesz spaces (vector lattices), such as band and the disjoint complement D^d of a subset D of a Riesz space. Let (Y, Σ, ν) be a σ-finite measure space. By $L_0(Y; \nu)$ we denote the space of all ν-measurable real functions on Y which are finite ν-a.e., with identification of functions which are equal ν-a.e. Let L be an ideal in $L_0(Y, \nu)$, i.e., L is a linear subspace with the additional property that $|g|\leqq|f|$, $g\in L_0(Y, \nu)$ and $f\in L$ implies that $g\in L$. A subset F of Y is called an L-null set if every $f\in L$ vanishes on F ν-a.e. There exists a maximal

Received January 5, 1984.

L-null set F_0, which is unique modulo null sets. The set $Y \setminus F_0$ is called the carrier of L. For the investigation of the ideal L we may assume that the carrier of L is equal to Y. We denote by $L_n^\sim$ the ideal in $L_0(Y, \nu)$ consisting of all functions g satisfying $\int_Y |fg|\, d\nu < \infty$ for all $f \in L$. The elements of $L_n^\sim$ can be identified in the obvious way with the linear functionals on L referred to in the introduction, which are called normal integrals. Let Y_0 be the carrier of $L_n^\sim$. In general Y_0 is a proper subset of Y. By way of example, if Y does not contain any atoms and $L = L_0(Y, \nu)$, then $Y_0 = \emptyset$ ([3], Example 85.1). Furthermore, $Y_0 = Y$ if and only if $L_n^\sim$ separates the points of L ([3], Theorem 95.2).

Now let (X, Λ, μ) be another σ-finite measure space and let M be an ideal in $L_0(X, \mu)$. The linear operator T from L into M is called an integral operator (or kernel operator) if there exists a $\mu \times \nu$-measurable function $T(x, y)$ on $X \times Y$, the kernel of T, such that for every $f \in L$,

$$(Tf)(x) = \int_Y T(x, y) f(y)\, d\nu(y) \quad \mu\text{-a.e. on } X.$$

Furthermore, T is called an absolute integral operator if the kernel $|T(x, y)|$ defines an integral operator from L into M as well. In fact, the integral operator T is absolute if and only if T is order bounded (i.e., T maps order intervals into order intervals), and in that case the absolute value $|T|$ of T in the Riesz space $\mathscr{L}_b(L, M)$ of all order bounded linear operators from L into M, is the integral operator with kernel $|T(x, y)|$ ([3], section 93). The set $\mathscr{I}(L, M)$ of all absolute integral operators from L into M is a band in $\mathscr{L}_b(L, M)$ ([3], Theorem 94.5). Observe that any integral operator T from L into M is order bounded as an operator from L into $L_0(X, \mu)$, since the kernel $|T(x, y)|$ defines an integral operator from L into $L_0(X, \mu)$.

As usual, for any $g \in L_n^\sim$ and $h \in M$ we denote by $g \otimes h$ the integral operator with kernel $h(x)g(y)$, and by $L_n^\sim \otimes M$ we denote the collection of all finite linear combinations of such operators. The elements of $L_n^\sim \otimes M$ are called integral operators of finite rank. It follows from $L_n^\sim \otimes M \subset \mathscr{I}(L, M)$ that the band $\{L_n^\sim \otimes M\}^{dd}$ generated by $L_n^\sim \otimes M$ satisfies $\{L_n^\sim \otimes M\}^{dd} \subset \mathscr{I}(L, M)$. In the next section we will show that $(L_n^\sim \otimes M\}^{dd} = \mathscr{I}(L, M)$, without any extra assumption about the carrier of $L_n^\sim$.

2. Let (Y, Σ, ν) be a σ-finite measure space and L an ideal in $L_0(Y, \nu)$. It will be assumed that the carrier of L is equal to Y. We begin with a lemma which characterizes ideals L for which $L_n^\sim = \{0\}$.

Lemma 2.1. *The following statements are equivalent.*

(i) $L_n^\sim = \{0\}$.

(ii) *For any* $A \in \Sigma$ *with* $0 < \nu(A) < \infty$ *there exist disjoint sets* $\{A_n\}_{n=0}^\infty$ *in* Σ *such that* $A = \bigcup_{n=0}^\infty A_n$, $\sum_{n=1}^\infty n\nu(A_n) = \infty$ *and* $\sum_{n=1}^\infty n\chi_{A_n} \in L$.

Proof. First observe that $L_n^\sim=\{0\}$ if and only if for every $A\in\Sigma$ with $0<\nu(A)<\infty$ there exists $0\leqq f\in L$ such that $\int_A f\,d\nu=\infty$. Now it is clear that (ii) implies (i), by taking $f=\sum_{n=1}^{\infty} n\chi_{A_n}$. Now assume that $L_n^\sim=\{0\}$ and let $A\in\Sigma$ with $0<\nu(A)<\infty$ be given. By the remark above, there exists $0\leqq f\in L$ such that $\int_A f\,d\nu=\infty$. Define

$$A_n=\{y\in A\colon\, n\leqq f(y)<n+1\}\quad (n=0,1,2,\ldots).$$

Then $\{A_n\}_{n=0}^{\infty}$ is a disjoint sequence and $\bigcup_{n=0}^{\infty} A_n=A$ (modulo a nullset). Moreover, since $\nu(A)<\infty$, it follows easily that $\sum_{n=1}^{\infty} n\nu(A_n)=\infty$, and it follows from $0\leqq\sum_{n=1}^{\infty} n\chi_{A_n}\leqq f$ that $\sum_{n=1}^{\infty} n\chi_{A_n}\in L$.

In the next proposition it is shown that an ideal L with $L_n^\sim=\{0\}$ cannot be the domain of any non-zero integral operator.

Proposition 2.2. *Let L be an ideal in $L_0(Y,\nu)$ with $L_n^\sim=\{0\}$, and let (X,Λ,μ) be a σ-finite measure space. If T is an integral operator from L into $L_0(X,\mu)$, then $T=0$.*

Proof. Assume that T is a non-zero integral operator from L into $L_0(X,\mu)$ with kernel $T(x,y)$. Since T is order bounded (and hence T is the difference of two positive operators), we may assume that $T>0$. Furthermore, since X and Y are both σ-finite, there exist $X'\in\Lambda$ and $Y'\in\Sigma$ with $0<\mu(X'),\ \nu(Y')<\infty$ such that $T'(x,y)=T(x,y)\chi_{X'}(x)\chi_{Y'}(y)$ is not equal to zero $\mu\times\nu$-a.e. on $X\times Y$. Let L' be the ideal in $L_0(Y',\nu)$ consisting of all restrictions of elements in L to Y'. Clearly $(L')_n^\sim=\{0\}$ and $T'(x,y)$ defines a non-zero integral operator from L' into $L_0(X',\mu)$. Hence, we may assume that (X,Λ,μ) and (Y,Σ,ν) are both finite measure spaces with $\mu(X)=\nu(Y)=1$. Furthermore, there exists $\varepsilon>0$ such that the set $P=\{(x,y)\in X\times Y\colon T(x,y)>\varepsilon\}$ satisfies $(\mu\times\nu)(P)>0$. Now it follows from $T(x,y)\geqq\varepsilon\,\chi_P(x,y)$ that the kernel $\chi_P(x,y)$ defines a non-zero integral operator from L into $L_0(X,\mu)$, Therefore, we may assume without loss of generality that $\mu(X)=\nu(Y)=1$ and the integral operator T from L into $L_0(X,\mu)$ has kernel $\chi_P(x,y)$ with $P\subseteqq X\times Y$ and $(\mu\times\nu)(P)=\delta>0$. It follows from

$$\int_Y\Big\{\int_X \chi_P(x,y)\,d\mu(x)\Big\}\,d\nu(y)=\delta>0,$$

and from $\nu(Y)=1$ that the set $A=\Big\{y\in Y\colon \int_X\chi_P(x,y)\,d\mu(x)\geqq\delta/2\Big\}$ satisfies $\nu(A)>0$.

Observe that if $B \subseteq A$, $B \in \Sigma$, then

$$\int_{X \times B} \chi_P \, d(\mu \times \nu) \geqq (\delta/2)\nu(B).$$

Since $L_n^{\sim} = \{0\}$, we can apply Lemma 2.1 to the set A, and so there exist disjoint sets $\{A_n\}_{n=0}^{\infty}$ in Σ with $\bigcup_{n=0}^{\infty} A_n = A$, $\sum_{n=1}^{\infty} n\nu(A_n) = \infty$ and $f = \sum_{n=1}^{\infty} n\chi_{A_n} \in L$. Put $g = Tf$. Since $\sum_{n=1}^{k} n\chi_{A_n} \uparrow_k f$, it follows from the theorem on integration of increasing sequences that $g = \sum_{n=1}^{\infty} nT\chi_{A_n}$ (μ-a.e. convergent series in $L_0(X, \mu)$). For $k = 1, 2, \ldots$ let $E_k = \{x \in X: g(x) \leqq k\}$. Then $E_k \uparrow X$. Since $\mu(X) < \infty$, there exists k such that $\mu(X \setminus E_k) < \delta/4$. Let this k be fixed.

Since $A_n \subseteq A$ ($n = 1, 2, \ldots$), it follows from the observation above that

$$(\delta/2)\nu(A_n) \leqq \int_{X \times A_n} \chi_P \, d(\mu \times \nu) = \int_{(X \setminus E_k) \times A_n} \chi_P \, d(\mu \times \nu) + \int_{E_k \times A_n} \chi_P \, d(\mu \times \nu) \leqq$$

$$\leqq \mu(X \setminus E_k)\nu(A_n) + \int_{E_k} \Big\{ \int_{Y} \chi_P(x, y)\chi_{A_n}(y) \, d\nu(y) \Big\} \, d\mu(x) \leqq$$

$$\leqq (\delta/4)\nu(A_n) + \int_{E_k} (T\chi_{A_n})(x) \, d\mu(x) \quad (n = 1, 2, \ldots),$$

and hence

$$(\delta/4)\nu(A_n) \leqq \int_{E_k} (T\chi_{A_n})(x) \, d\mu(x) \quad (n = 1, 2, \ldots).$$

This implies that

$$(\delta/4) \sum_{n=1}^{\infty} n\nu(A_n) \leqq \int_{E_k} \Big\{ \sum_{n=1}^{\infty} n(T\chi_{A_n})(x) \Big\} \, d\mu(x) = \int_{E_k} g(x) \, d\mu(x) \leqq k\mu(E_k) < \infty,$$

which is a contradiction. This completes the proof of the proposition.

Corollary 2.3. *Let (X, Λ, μ) and (Y, Σ, ν) be σ-finite measure spaces and let L and M be ideals in $L_0(Y, \nu)$ and $L_0(X, \mu)$ respectively. Denote by Y_0 the carrier of $L_n^{\sim}$ and by X_0 the carrier of M. Let T be an integral operator from L into M with kernel $T(x, y)$. Then $T(x, y) = 0$ $\mu \times \nu$-a.e. outside $X_0 \times Y_0$.*

Proof. Let K denote the ideal in $L_0(Y \setminus Y_0, \nu)$ consisting of all restrictions of elements in L to $Y \setminus Y_0$. It is clear that $K_n^{\sim} = \{0\}$. Let $S(x, y)$ be the restriction of $T(x, y)$ to $X \times (Y \setminus Y_0)$. Then the kernel $S(x, y)$ defines an integral operator S from K into M. It follows now from the above proposition that $S = 0$, and hence $S(x, y) = 0$ $\mu \times \nu$-a.e. on $X \times (Y \setminus Y_0)$ (see [3], Theorem 93.1). Therefore $T(x, y) = 0$ $\mu \times \nu$-a.e. on $X \times (Y \setminus Y_0)$. Furthermore, the integral operator from L into M with kernel $T(x,y)\chi_{X\setminus X_0}(x)$ is the zero operator, hence $T(x,y)\chi_{X\setminus X_0}(x)=0$ $\mu\times\nu$-a.e. on $X\times Y$, i.e., $T(x,y)=0$ $\mu\times\nu$-a.e. on $(X\setminus X_0)\times Y$. We may conclude, therefore, that $T(x,y)=0$ $\mu\times\nu$-a.e. outside $X_0\times Y_0$.

In the next theorem we will show that the band of absolute integral operators from the ideal L into the ideal M is equal to the band generated by the integral operators of finite rank. Under the additional assumption that the carrier of the ideal $L_n^\sim$ is equal to the whole space Y this result was proved by A. V. Buhvalov [1] (see also [3], Theorem 95.1).

Theorem 2.4. *Let* (Y, Σ, ν) *and* (X, Λ, μ) *be σ-finite measure spaces and* L *and* M *ideals in* $L_0(Y,\nu)$ *and* $L_0(X,\mu)$ *respectively. Then the band* $\mathscr{I}(L,M)$ *of all absolute integral operators from* L *into* M *is equal to the band* $\{L_n^\sim \otimes M\}^{dd}$.

Proof. It is sufficient to show that any positive integral operator T from L into M is contained in $\{L_n^\sim \otimes M\}^{dd}$. Let $T(x,y)$ be the kernel of T. Denote by Y_0 the carrier of $L_n^\sim$ and by X_0 the carrier of M. By the above corollary we have $T(x,y)=0$ $\mu\times\nu$-a.e. on $(X\times Y)\setminus(X_0\times Y_0)$. Since Y_0 is the carrier of $L_n^\sim$, there exists a sequence $\{Y_n\}_{n=1}^\infty$ in Σ such that $Y_n\uparrow Y_0$ and $\chi_{Y_n}\in L_n^\sim$ for all n. Similarly, there exists a sequence $\{X_n\}_{n=1}^\infty$ in Λ with $X_n\uparrow X_0$ and $\chi_{X_n}\in M$ $(n=1,2,\ldots)$. Now define

$$T_n = \inf\{T, n\chi_{Y_n}\otimes\chi_{X_n}\} \quad (n=1,2,\ldots),$$

which is an integral operator with kernel

$$T_n(x,y) = \inf\{T(x,y),\ n\chi_{X_n}(x)\chi_{Y_n}(y)\}$$

(see [3], Theorem 94.3). Note that $0\leqq T_n\in\{L_n^\sim \otimes M\}^{dd}$. Since $T(x,y)=0$ $\mu\times\nu$-a.e. outside $X_0\times Y_0$, it follows that $0\leqq T_n(x,y)\uparrow T(x,y)$ $\mu\times\nu$-a.e. on $X\times Y$, and hence $0\leqq T_n\uparrow T$ in $\mathscr{L}_b(L,M)$ ([3], Theorem 94.5). This shows that $T\in\{L_n^\sim \otimes M\}^{dd}$.

Remark. 2.5. In [1] Buhvalov presented an important characterization of integral operators. Recall that the sequence $\{f_n\}_{n=1}^\infty$ in $L_0(Y,\nu)$ is called star convergent to zero (denoted by $f_n\xrightarrow{*}0$) if every subsequence of $\{f_n\}_{n=1}^\infty$ contains a subsequence which is converging to zero ν-a.e. on Y. Let L and M be ideals in $L_0(Y,\nu)$ and $L_0(X,\mu)$ respectively. We say that the operator T from L into M satisfies Buhvalov's condition if it follows from $0\leqq u_n\leqq u\in L$ $(n=1,2,\ldots)$ with $u_n\xrightarrow{*}0$ that $Tu_n\to 0$ pointwise ν-a.e. on X. It was proved by Buhvalov that the operator T from L into M is an integral operator if and only if T satisfies Buhvalov's condition. In the proof of this theorem it is assumed first that the carrier of $L_n^\sim$ is equal to Y (see also [3], Theorems 96.5 and 96.8), and then it is observed that the general case can be reduced easily to this special situation. Indeed, consider the ideal $K=L\cap L_\infty(Y; \nu)$ and let T_0 be the restriction of T to K. If T satisfies Buhvalov's condition, then T_0 satisfies Buhvalov's condition as well and the carrier of $K_n^\sim$ is equal to Y. Hence T_0 is an integral operator with kernel $T(x, y)$. Now it is easy to see that T is an integral operator with kernel $T(x, y)$.

References

[1] A. V. BUHVALOV, Integral representation of linear operators, *Zapiski Nauchnykh Seminarov Leningradskogo Otdeleniya Matematicheskogo Instituta im. V. A. Steklova AN SSSR*, 47 (1974), 5—14; English translation in: *Soviet Math.*, 9 (1978), 129—137.

[2] W. A. J. LUXEMBURG and A. C. ZAANEN, *Riesz Spaces. I*, North-Holland (Amsterdam—London, 1971).

[3] A. C. ZAANEN, *Riesz Spaces. II*, North-Holland (Amsterdam—New York—Oxford, 1983).

DEPARTMENT OF MATHEMATICS
DELFT UNIVERSITY OF TECHNOLOGY
JULIANALAAN 132
2628 BL DELFT, THE NETHERLANDS

Bibliographie

Advances in Probability Theory: Limit Theorems and Related Topics, Edited by A. A. Borovkov (Translation Series in Mathematics and Engineering), XIV+377 pages, Optimization Software, Inc., Publication Division, New York, 1984. (Distributed by Springer-Verlag.)

This is the English translation of a fine collection of seventeen strong papers on limit theorems and some applications written in the best tradition of the Soviet probability school. The collection was originally published in Russian as Volume 1 of the Proceedings of the Institute of Mathematics, Siberian Branch of the USSR Academy of Sciences, Nauka, Novosibirsk, 1982.

The leading themes are the rate of convergence in the invariance principle and the theory of large deviations, applications include hypothesis testing, branching processes, stochastic differential equations, martingales in the plane, quadratic variation, queueing theory and some other fields. The authors are K. Arndt, I. S. Borisov, A. A. Borovkov, V. M. Borodihin, V. I. Chebotarev, S. G. Foss, G. P. Karev, V. I. Lotov, A. A. Mogul'skij, S. V. Nagaev, G. V. Nedogibchenko, I. F. Pinelis, A. I. Sahanenko, L. Ya. Savel'ev, V. A. Topchij, and V. V. Yurinskij.

Let us hope that the similarly strong subsequent volumes of the Novosibirsk Proceedings will be translated as well, and soon.

Sándor Csörgő (Szeged)

Asymptotic Analysis II-Surveys and New Trends, Edited by F. Verhulst (Lecture Notes in Mathematics, 985), III+497 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1983.

The methods of asymptotic analysis are widely applied in physics, engineering sciences, biology, etc. Research papers on such applications are continually being published in various journals, but collections like the present one are necessary to enable specialists to see what goes on in a broader context. The predecessor of the present volume is *Asymptotic analysis. From theory to applications* (Lectures Notes in Mathematics, 711) published four years ago. A somewhat similar good collection is *Theory and Applications of Singular Perturbations* (Lecture Notes in Mathematics, 942) though the main feature of these proceedings is different from that of the present one and its predecessor.

This volume is divided into three parts: Survey Papers, Survey Papers with Research Aspects, and Research Papers. The list of the authors reads as follows: Z. Schuss and B. J. Matkowsky, V. F. Butuzov and A. B. Vasil'eva, J. Grasman, A. S. Bakaj, F. Verhulst, E. Sanchez-Palencia, V. Drăgan and A. Halanay; Robert E. O'Malley, Jr., P. W. Hemker, Jan A. Sanders, Richard Cushman, A. van Harten; G. G. Rafel, V. Drăgan, J. Grasman adn B. J. Matkowsky, A. H. P. van der Burgh, V. N. Bogaevsky and A. Ya. Povzner, Wiktor Eckhaus.

The main feature is the application of asymptotics to problems concerning certain differential equations, ordinary or partial, and systems motivated by practical problems and depending on parameters. Also many qualitative results are given. The papers in Part 3 "... are settling old questions in the literature and are opening up new lines of research both in the field of methods in applied mathematics and its applications".

J. Hegedűs (Szeged)

J. P. Aubin—A. Cellina, Differential Inclusions (Grundlehren der mathematischen Wissenschaften, 264), XIII+342 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984.

The differential equation $x'=f(t, x)$ as a model of an evolving system expresses that at every moment the evolution of the system is uniquely determined by its instantaneous state. The existence and uniqueness theorems say that the future (or even the past) of a system described by a differential equation model is uniquely determined by its state at a fixed moment. This cruelly deterministic model can be quite convenient for describing systems that arise in physics, mechanics, engineering if one ignores the influence of the environment on the system. But if one wants to describe macro-systems with mutual effects, some of which can be only estimated, one has to take into account the uncertainty in the model. This uncertainty may be caused by the absence of controls, but it can appear even in a controlled system as a consequence of the ignorance of the laws relating the controls and the states of the system. In such cases the instantaneous state of the system does not determine uniquely the velocity of its change. Mathematically, to every moment t and every state x of the system a set-valued map $F=F(t, x)$ associates to t and x the set of feasible velocities, i.e. $x' \in F(t, x)$.

This new model is perceptibly more complicated and raises new mathematical problems. Even in the simplest cases one needs a pretty wide knowledge from topology and functional analysis. Fortunately, in the two first chapters the authors give a good review on the prerequisite and an excellent introduction to the theory of set-valued maps. The existence problems are treated by the same methods (approximation, fixed-point techniques) as in the theory of differential equations but very interesting new problems arise. For example, while for ordinary differential equations th convergence of the approximate solutions implies the convergence of their derivatives, this is not the case any more for differential inclusions. In fact the convergence of the derivatives will be the main issue in each existence proof. A separate chapter is devoted to the existence and qualitative properties of solutions of differential inclusions with maximal monotone maps.

One can imagine that the set of trajectories is large in this new model, so it is very important in the theory to find a devising mechanism for selecting special trajectories. Such a mechanism is given by Optimal Contral Theory: it selects the trajectories that optimize a given criterion, a functional on the space of all trajectories. The method requires the existence of a decision maker who controls the system having a perfect knowledge of the future and makes a choice once and for all at the origin of the period of time. These requirements are not met by the interactive systems evolving, e.g., according to the laws of Darwinian evolution. They do not optimize any criterion, simply they face a minimal requirement (called viability), they must meet to remain alive. Mathematically, this means that the trajectories have to remain in a set of constraints. The viability theory guarantees the existence of such trajectories. As an application, a dynamical analog is established to the Walras equilibrium theory on the price decentralization in economy. Here the viability constraint is the requirement that the sum of the consumed commodity boundles lies in the set of available goods.

Stability theory for differential inclusions based on Lyapunov functions concludes the book.

This well-written excellent monograph summarizes the new methods and results in this very important modern field. It can be highly recommended to mathematicians and users of mathematics interested in dynamical systems and applications.

L. Hatvani (Szeged)

Michael J. Beeson, Foundations of Constructive Mathematics, Metamathematical Studies (Ergebnisse der Mathematik und ihrer Grenzgebiete, 3. Folge, Band 6) XIII+466 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1985.

The volume, devoted to constructive mathematics, consists of four parts and a historical survey. The first part, entitled "Practice and Philosophy of Constructive Mathematics", is intended to give an overview of some principal ideas on both the philosophy and the foundations of explicite existential proofs. In particular, basic methods and results important on their own right as well as for further usage are collected in this first part. The title of the second part is "Formal Systems of the Seventies". Central to the discussions in this part are three systems: theories of rules, sets and classifications proposed by Feferman, type theories due to Martin-Löf and constructive set theories of Myhill and Friedmann. The third part is devoted, to "Metamathematical Studies". In patricular, Aczel's iterative sets are investigated in detail, and, also, questions related to recursiveness, continuity and uniform continuity are considered. The final part ("Metaphilosophical Studies") is a continuation of the work started by Kreisel and Goodman in developing the "theories of constructions" which serve as the foundation to philosophical discussions on the notion of "constructive proof". The volume ends with a long survey of the historical development of the subject, emphasizing the contributions of the two outstanding forefathers Bishop and Martin-Löf.

This clearly written and attractive book is an exciting and useful reading to most mathematicians interested in the theory of algorithms, proofs and foundations of mathematics and computer sciences.

P. Ecsedi-Tóth (Szeged)

G. W. Bluman, Problem Book for First Year Calculus, (Problem Books in Mathematics), XV+385 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1984.

Here is a new excellent book in the series "Problem Books in Mathematics". It contains more than 1000 problems, 300 of which are solved in great detail.

In Chapters I and II we find problems in graphing and geometry: Chapters II, IV, V and VI, respectively, contain problems belonging to physics and engineering, business and economics, biology and chemistry, and numerical methods. Finally, Chapters VII and VIII provide problems on the theory and techniques of calculus.

The organization of the first six chapters is the same. After a short discussion of background material there are several thoroughly solved problems and then proposed problems follow. (Answers to these problems can be found in Chapter IX.)

There are routine and difficult problems, almost all of which are closely connected with practical questions.

Let us mention a few characteristic examples. 1) A sprinter who runs the 100 metre in 10.2 seconds accelerates at a constant rate for the first 25 metres and then continues at a constant speed for the rest of the race. Find his acceleration. 2) A radar antenna is mounted 500 metres horizontally away from, and is aimed at, a rocket sitting on a launching pad. The rocker blasts off at time $t=0$ and thereafter climbs vertically with a constant acceleration of 10 m/s^2. If the antenna remains aimed directly at the rocket, how fast must it be rotating upward 10 sec after blast-off? 3) What should the speed limit be for cars on the Lions Gate Bridge in Vancouver, British Columbia, during rush hour traffic, in order to maximize the flow of traffic? (This is of course an "open-ended" problem. The answer depends on one's assumptions. There are models in the book.) 4) A manufacturer estimates that he can sell 2000 toys per month if he sets the unit price at $ 5.00. Furthermore he estimates that for each $ 0.20 decrease in price his sales will increase by 200 per month. a) Find the demand and revenue functions. b) Find the number of toys that he should sell each month in order to maximize the monthly revenue. c) What is the maximum monthly revenue? 5) A contagious disease, say smallpox, begins to spread in a community of 1000 people. This disease has the property that it spreads by contact and that each person who has it immediately and forever infects others. Initially, one person has it, and the speed of the epidemic appears to lessen after a month. Find how many people have had the disease at any time.

All in all this is a well thought out, stylish book, which one could safely put into the hands of future users of mathematics.

L. Pintér (Szeged)

K. W. Chang—F. A. Howes, Nonlinear Singular Perturbation Phenomena: Theory and Applications, (Applied Mathematical Sciences, 56), VIII+180 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1984.

In applications one often meets differential equation models containing a "small parameter" ε. The equation belonging to $\varepsilon=0$ is called the unperturbed equation and the original one is called the perturbed equation. If the perturbed equation in normal form depends on the small parameter regularly (continuously, smoothly) then the perturbation is called regular, otherwise — singular. In the regular case the solutions of the unperturbed equation are good approximations, but in the singular case the small terms may be neglected only if certain special conditions are met. These are determined by the theory of singular perturbations.

In this book the authors are dealing with those singular perturbations when the basic (scalar or vector) equation is of second order and the second derivative of the unknown function is multiplied by a small parameter.

The best course to characterize the book is to cite some sentences from the authors' preface: "Our purpose in writing this monograph is twofold. On the one hand, we want to collect in one place many of the recent results on the existence and asymptotic behavior of solutions of certain classes of singularly perturbed nonlinear boundary value problems. On the other, we hope to raise along the way a number of questions for further study, mostly questions we ourselves are unable to answer. ... We offer our results with some trapidation, in the hope that they may stimulate further work in the challenging and important area of differential equations."

To the problems treated here the well-known methods, such as the methods of matched asymptotic expansions and two-variable expansions are not immediately applicable, differential inequality techniques coupled with geometric and asymptotic concepts are used instead. The book is concluded by very interesting examples and applications (e.g. equations from theory of nonpremixed combustion, catalytic reaction theory).

This monograph will be very useful not only for experts in singular perturbations, who get a good survey on results that were available before only from articles, but for mathematicians and science students interested in differential equations and their applications.

L. Hatvani (Szeged)

Convex analysis and optimization, Edited by J.-P. Aubin, R. B. Vinter (Research Notes in Mathematics, 57), VIII+210 pages. Pitman Advanced Publishing Program, Boston—London —Melbourne, 1982.

This book consists of 8 papers devoted to surveying new results in convex analysis and its applications in optimization theory. The papers by J. P. Aubin, J. Ekeland and A. D. Ioffe concern nonsmooth analysis. The papers by L. C. Young and R. B. Vinter illustrate the role of convexity in optimization theory and modelling problems even in cases which are not convex in a conventional sense. In their paper J. E. Jayne and C. A. Rogers investigate measurable selections of multivalued mappings. The papers by J. B. Hiriart-Urruty and J. P. Crouzeix treat a generalization of the notion of the subdifferential of a convex function.

László Gehér (Szeged)

John B. Conway, A Course in Functional Analysis (Graduate Texts in Mathematics, 96), XIV+404 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1984.

This is an excellent textbook on functional analysis. The author's main pedagogic purpose seems to be to help the students acquire an intuition in the subject. According to this purpose, often the book does not discuss immediately the greatest generality possible, but it starts from the particular and works its way to the more general case. The contents are arranged subject to this principle.

Chapter I introduces the geometry of Hilbert space. Chapter II begins the discussion of operators on a Hilbert space. This chapter contains the spectral theory and functional calculus for compact normal operators. In Chapter III Banach spaces are studied. Among others, the Hahn—Banach Theorem, the Open Mapping Theorem and the Principle of Uniform Boundedness are proved. Chapter IV deals with locally convex spaces and Chapter V treats the weak and weak-star topologies. In Chapter VI bounded operators on a Banach space are discussed. The subject of Chapter VII is the discussion of Banach algebras and the spectral theory for operators on a Banach space. Chapter VIII deals with C^*-algebras. Chapter IX presents the Spectral Theorem and its ramification in the framework of C^*-algebras. This chapter is concluded by the multiplicity theory of normal operators. Chapter X deals with unbounded operators in a Hilbert space and Chapter XI is devoted to the Fredholm theory.

Some of the prerequisities are presented in the appendicies. Appendix A contains, among others, a discussion of nets, Appendix B deals with the dual of $L^p(\mu)$ and Appendix C discusses the dual of $C_0(X)$. On the last pages, a rich bibliography is found, a list of symbols is given, and an index completes the book.

A great number of examples helps understand the abstract concepts, makes the presentation colorful, and enlightens the connection with some other branches of mathematics. The book also contains a great number of exercises of varying degrees of difficulty. Some of them are routine, some demand proofs left to the reader in the text, others extend the theory. By all means, the method of the discussion stimulates the reader not only to read but to do mathematics.

The prerequisities for this book are a firm foundation in measure and integration theory and some knowledge of point set topology. Analytic functions are used to furnish some examples and, in the second half of the book, analytic function theory is also applied in the proofs of results.

This book is warmly recommended first of all to graduate students. It contains a fairly large amount of material. If the reader has no time or energy to go through the whole book, he can leave out the sections marked with an asterisk. It could also serve as a handbook for a researcher in functional analysis.

E. Durszt (Szeged)

Jonh B. Conway, Functions of One Complex Variable, Second Edition (Graduate Texts in Mathematics, 11), XI+317 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1984.

This second edition contains some substantial revisions. These are: the inclusion of John Dixon's treatment of Cauchy's Theorem (keeping the original homotopic version), a new, elementary proof of Runge's Theorem due to Sandy Grabiner, a new appendix containing a guide for further reading and several additional exercises. These changes make this beautiful, succesful work even more exciting. It is written in a very clear style, its reading requires only the knowledge of very basic facts from calculus. This delicious, effective introductory book can be warmly recommended to every student who wants to get acquainted with the classical field of complex analysis.

L. Kérchy (Szeged)

Joseph Diestel, Sequences and Series in Banach Spaces (Graduate Texts in Mathematics, 92), XI+261 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1984.

This volume has been written to acquaint a wide audience of mathematicians with the most important results and methods of the theory of Banach spaces. The author writes in the introduction: "I have concentrated on presenting what I believe are basic phenomena in Banach spaces that any analyst can appreciate, enjoy, and perhaps even use." The technical jargon of the area is avoided, the style is very informal, the text holds the reader. The selected material is arranged into the following chapters: Introduction. — Riesz's Lemma and Compactness in Banach Spaces.—The Weak and Weak* Topologies: an Introduction. — The Eberlien — Šmulian Theorem. — The Orlicz — Pettis Theorem. — Basic Sequences. — The Dvoretsky — Rogers Theorem. — The Claasical Banach Spaces.—Weak Convergence and Unconditionally Convergent Series in Uniformly Convex Spaces.— Extremal Tests for Weak Convegence of Sequences and Series. — Grothendieck's Inequality and the Grothendieck — Lindenstrauss — Pelczynski Cycle of Ideas. An Intermission: Ramsey's Theorem. — Rosenthal's l_1-theorem. — The Josefson — Nissenzweig Theorem. — Banach Spaces with Weak* — Sequentially Compact Dual Balls. — The Elton — Odell $(1+\varepsilon)$-Separation Theorem.

At the end of every chapter several exercises, historical remarks and a detailed bibliography can be found.

L. Kérchy (Szeged)

Differential Geometry and Complex Analysis. A volume dedicated to the memory of Harry Ernest Rauch, Edited by I. Chavel and H. M. Farkas, XIII+222 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1985.

This book is dedicated to the memory of the excellent mathematician Harry Ernest Rauch (1925—1979) who made significant contributions to the areas of pinching theorems of Riemannian manifolds, Teichmüller theory and theta function theory. The volume contains his biography, bibliography and the list of Ph. D. theses written under his supervision. The influence of his work on the progress of differential geometry, complex analysis and theta function theory is presented in the survey articles written by M. Berger, C. J. Earle and H. M. Farkas. The other papers published in this volume give the reader ideas about the new development of the fields of mathematics which are connected with Rauch's researches and interests. 7 papers are devoted to the study of various questions of differential geometry (by E. Calabi, J. Cheeger, M. Gromov, S. S. Chern, D. Gromoll, K. Gove, W. Klingenberg, A. Marden, K. Strebel and S. T. Yau), 3 papers to geometric group theory (L. V. Ahlfors, I. Kra, Min-Oo and E. A. Ruh), 2 papers to complex function theory (J. M. Anderson, F. W. Gehring, A. Hinkkanen and L. Bers) and 1-1 papers to Teichmüller theory (by R. D. M. Accola) and to elliptic differential operators (by L. Nirenberg).

The book is warmly recommended to everybody working in differential geometry and complex function theory.

Péter T. Nagy (Szeged)

Harold M. Edwards, Galois Theory (Graduate Texts in in Mathematics Vol. 101) XII+152 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1984.

This book is a constructive and concrete introduction to Galois theory presenting an exposition of the theory in tems near enough to Galois' original work, the "Mémoire sur les conditions de résolubilité des equations par radicaux" which is contained as Appendix 1 in this book in an English translation due to the author.

It is devided into eight parts, each of which is followed by a set of exercises.

After a brief historical outline of the solutions of quadratic, cubic and quartic equations from 1700 B. C. to 1600 A. D., the Lagrange (Vandermonde) resolvent is introduced and its application in solving equations of degree 2, 3 and 4 is discussed briefly.

A subsequent part is devoted to the cyclotomic equations and to some results of Gauss on constructing p-gons by ruler and compass.

The remaining five parts can be considered, with one major exception, as a detailed version of Galois' original work, which contains full and rigorous proofs (that are missing at Galois). The exception mentioned is the material on factorization of polynomials due to Kronecker. This gives clear meaning to the computations with roots of algebraic equations that Galois and Lagrange performed without inhibition and comment.

Appendices 2 and 3 help the reader understand the original text of Galois making connections with the full (and modern) treatment of the problem contained in the book.

This interesting volume ends with the complete solutions of the exercises.

Lajos Klukovits (Szeged)

Leonard Euler, Elements of Algebra, LX+593 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1984.

The eighteenth century was, par excellence, a textbook age in mathematics, never before had so many books appeared in so many editions. Euler also composed a popular algebra textbook, the famous "Volstandige Anleitung zur Algebra", which appeared at St. Petersburg in 1770. The exceptional didactic quality of this book is attributed to the fact that it was dictated by the blind author through a relatively untutored domestic. In 1774 appeared the French version of this book, translated by M. Bernoulli, with Additions composed by Lagrange.

The present volume is a reprint of the fifth English edition which appeared in 1840. This version is a translation, due to J. Hewlett, of the French edition, which was prefixed by "A Memoire of the Life and Character of Euler" by Francis Horner. It also contains a paper of C. Truesdell under the title "Leonard Euler, Supreme Geometer" which was originally published by the University of Wisconsin Press in 1972.

This book, which summarises the eighteenth century knowledge on determinate and indeterminate quantities and equations, illustrates — as other algebra textbooks of the century — a tendency toward increasingly algorithmic emphasis, while at the same time there remained considerable uncertanity about the logical bases of the subject.

We warmly recommend this book to everybody who wants to read "The Masters".

Lajos Klukovits (Szeged)

L. R. Foulds, Combinatorial Optimization for Undergraduates, XII+228 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1984.

This is a book on combinatorial optimization covering the main topics in this field and written especially for undergraduates. It is a good introductory book in which clarity and plausability have priority over rigorous arguments. Each topic begins with solving a concrete numerical example, the general problem and several algorithms come only afterwards. Some exercises are also provided. Sometimes, when it serves a better understanding of an algorithm, rigorous proofs also occur.

The first two chapters deal with linear programming (simplex method, transportation problem, assignment problem), integer programming and dynamic programming. Chapter 3 deals with the optimization on graphs and networks; minimal spanning trees, shortest paths, the maximum-flow problem, the minimum-cost-flow problem and activity networks are discussed. Chapter 4 handles various problems in operation research, engineering and biology; this chapter includes the travelling salesman problem, the vehicle scheduling problem, evolutionary trees, car pooling and facilities layout.

"The book is intended for undergraduates in mathematics, engineering, business, or the phisical or social sciences. It may also be useful as a reference text for practising engineers and scientists."

G. Czédli (Szeged)

Daniel S. Freed—Karen K. Uhlenbeck, Instantons and Four-Manifolds, (Mathematical Sciences Research Institute Publications, 1), 232 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1984.

This book is devoted to the systematic study of topological, differential geometric and nonlinear analytic methods used for the proof of Donaldson's and Freedman's theorems on the existence of exotic differentiable structures on four-manifolds. As originally Donaldson pointed out, these topological theorems are closely related to the properties of solution spaces of Yang-Mills and self-dual equations. The solution space of self-dual equations on a four-manifold M is divided out by a natural equivalence giving the "moduli space" $\mathfrak{M}$. This moduli space can be regarded as an oriented five-manifold with point singularities, the neighbourhoods of singular points are cones on the complex projective plane and M is the boundary of $\mathfrak{M}$. The geometric description of the moduli space can be applied to the investigation of exotic topologies of the basic manifold M.

The reader is assumed to be familiar with the theory of elliptic differential operators and the smoothing theory of manifolds.

Péter T. Nagy (Szeged)

J. K. Hale—L. T. Magalhães—W. M. Oliva, An Introduction to Infinite Dimensional Dynamical Systems-Geometric Theory (Applied Mathematical Sciences, Vol. 47), 195 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1984.

"The purpose of these notes is to outline an approach to the development of a theory of dynamical systems in infinite dimensions which is analogous to the theory of finite dimensions... the discussion centers around retarded functional differential equations although the techniques and several of the results apply to more general situations; in particular, to neutral functional differential equations, parabolic partial differential equations and some other types of partial differential equations" as the authors write in the introduction.

The book consists of eleven sections and an appendix. The first section comprises an abstract formulation of a class of dynamical systems and propounds some of the basic questions that should be discussed.

Sections 3 and 4 are concerned with retarded functional differential equations on manifolds. An existence and uniqueness theorem is stated and some important properties of the solution map are given.

For ordinary differential equations the Kupka—Smale theorem asserts that the property that all critical points and periodic orbits are hyperbolic and the stable and unstable manifolds intersect transversally is generic in the class of all ordinary differential equations in R^n. Although presently there is no complete proof of the Kupka—Smale theorem for retarded functional differential equations, some interesting results in this direction are proved in Section 4.

The attractivity properties of the invariant sets and limit sets are studied in Section 5.

The set of all initial data of global bounded solutions of a given retarded functional differential equation is called the attractor. The structure of the attractor (dimension, smoothness) is investigated in Sections 6 and 7.

Section 8 discusses the difficulties of the generalization (if it exists) of the Hartman—Grobman theorem for retarded functional differential equations.

Poincaré's compactification method is extended in Section 9 to get the behaviour at infinity of solutions of a linear delay equation.

Section 10 proves stability theorems for Morse—Smale maps.

The book is finished by bibliographical notes (Section 11) and an appendix written by K. P. Rybakowski entitled "An Introduction to the Homotopy Index Theory in Noncompact Spaces".

These notes give not only a unified exposition of the fundamental results in this field but also contain speculations on some of the possible directions for future research.

T. Krisztin (Szeged)

Infinite-Dimensional Systems, Proceedings, Retzhof, 1983. Edited by F. Kappel and W. Schappacher (Lecture Notes in Mathematics, 1076), VII+278 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984.

This volume contains lectures at the Conference on Operator Semigroups and Applications held in Retzhof, Austria, June 5—11, 1983. "The aim of this conference was to stimulate the exchange of ideas and methods and to provide information on recent advances in various directions of research" as the editors write in the introduction. This aim is completely fulfilled.

Most of the papers have the common feature that they use the so-called semigroup approach. This method can quite effectively be applied for the investigation of partial differential equations or partial functional differential equations when the original equation is rewritten as an ordinary or a functional differential equation in a suitable infinite dimensional Banach space. In addition, there are lectures taken from various branches of the theory of differential equations.

Since there is no room to list the titles of all of the 22 papers, here are only the main key words and phrases to show that this volume is important and interesting for everyone who is interested in the theory and applications of operator semigroups: generators of positive semigroups, Wiener's theorem and semigroups, nonlinear diffusion problems, abstract Volterra equations, stability by Lyapunov functionals, extrapolation spaces, wave propagation, Burgers systems, age-dependent population dynamics, semilinear periodic-parabolic problems, approximation of semigroups, differentiability of semigroups, semigroups generated by convolution equations, the Sharpe—Lotka theorem, integrable resolvent operators, functional evolution equations, rate of convergence in singular perturbations.

T. Krisztin (Szeged)

Dennis Kletzing, Structure and Representations of Q-Groups (Lecture Notes in Mathematics 1084), VI+290 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984.

A Q-group is a finite group all of whose ordinary complex representations have rationally valued characters. The representation of a group by means of permutations is always a basic question in the theory considered and the work of Frobenius and Young shows that the rationally represented characters of the symmetric and hyperoctahedral groups are generalized permutation characters. One can also mention Artin's theorem in this connection, asserting that a rationally represented character of a group may be written as a linear combination of permutation characters with rational numbers as coefficiens.

The book is a clear exposition and a further extension of this theory concentrating on the following two questions: I. What can be said about the structure of a Q-group? II. Under what circumstances can we conclude that the rationally represented characters of a Q-group are generalized permutation characters? Chapters 1 and 2 are devoted to question I, while Chapters 3 and 4 concentrate on II.

The most important invariant of this theory is the smallest positive integer $\gamma(G)$ with the property that $\gamma(G)\chi$ is a generalized permutation character of the group G whenever χ is a rationally represented character of G. Furthermore, the concepts of local classes, local characters and local splitting are basic in this consideration. The main result of the book is the following theorem proved in Chapter 4: If G is a Q-group which is locally split on every local class, the $\gamma(G)=1$. Several applications of these results to the symmetric groups and Weyl groups are considered.

The volume assumes that the reader is familar in group theory, commutative algebra and the ordinary representation theory of finite groups. It will be of interest to researches working in finite groups and their representation theory and also to algebraic geometers.

Z. I. Szabó (Szeged)

M. A. Krasnosel'skiĭ,—P. P. Zabreĭko, Geometrical Methods of Nonlinear Analysis (Grundlehren der mathematischen Wissenschaften, 263) XIX+409 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984.

The appearance of the computers has had an influence also on the development of the theory of nonlinear operator equations. The numerical methods have come by a very important role. However, there are many problems in the theory which are essential also from the point of view of practice, and the numerical approach to which is difficult or not possible at all. Such are the stability of solutions, bifurcation and branching of solutions, criteria for the existence of periodic solutions, the analysis of the structure of the solutions, etc. Even the numerical analysis raises such problems for the iteration processes and other procedures. Undoubtedly, the geometrical or topological methods are the most important and widely used tools for the study of these problems. The book, which is a translation of the Russian original edition from 1975, gives a systematic and complete treatment of these methods.

The central concept is an elemantary integer-valued invariant for vector fields called rotation. This tool makes it possible to investigate not only single operator equations, but general classes of operators which permit certain modifications of the equations, e.g., continuous deformation or perturbation of constituents of the equation. The first four chapters are devoted to computing and estimating the rotations of different vector fields and applications to periodic problems. Chapters 5 and 6 deal with the existence of the solutions and nontrivial solutions and with the estimation of the number of solutions for equations with nonlinear operators. Chapter 7 illuminates well how useful this approach can be in many topics of nonlinear analysis and qualitative theory (approximation methods for the solution of nonlinear operator equations and stability theory for instance). Among others, it is proved here that if a system has an isolated equilibrium state which is Lyapunov-stable, then, in general, it cannot be continuously deformed so that it reaches an equilibrium state which is asympotically stable. The last chapter deals with the influence of small perturbations.

The results are illustrated by applications to nonlinear vibrations, nonlinear mechanical problems, nonlinear integral equations and boundary value problems.

This book will be indispensable for mathematicians and other scientists interested in qualitative problems for operator equations.

L. Hatvani (Szeged)

Lie Group Representations III, Proceedings of the Special Year held at the University of Maryland, College Park 1982—1983. Edited by R. Herb, R. Johnson, R. Lipsman and J. Rosenberg (Lecture Notes in Mathematics, 1077), X+454 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984.

During the academic year 1982—1983, the Department of Mathematics of the University of Maryland conducted its traditional Special Year in the modern theory of Lie group representations. The articles by the invited speakers, distinguished researchers, have been published in a 3-volume set, the last of which is the present book.

The table of contents: 1. L. Corwin: Matrix coefficients of nilpotent Lie groups. 2. L. Corwin: Primary projections on nilmanifolds. 3. L. Corwin: Solvability of left invariant differential operators on nilpotent Lie groups. 4. M. Cowling and A. Korányi: Harmonic analysis on Heisenberg type groups from a geometric viewpoint. 5. M. Duflo: On the Plancherel formula for almost algebraic real Lie groups. 6. M. Felnsted-Jensen: Harmonic analysis on semisimple symmetric spaces. A method of duality. 7. B. Helffer: Partial differential equations on nilpotent groups. 8. S. Helgason: Wave equations on homogeneous spaces. 9. R. Howe, G. Ratcliff and N. Wildberger: Symbol mappings for certain nilpotent groups. 10. H. Moscovici: Lefschetz formulae for Hecke operators. 11. R. Penney: Harmonic analysis on unbounded homogeneous domains in C^n. 12. W. Rossmann: Characters as contour integrals. 13. L. Preiss Rothschild: Analyticity of solutions of partial differential equations on nilpotent Lie groups. 14. V. S. Varadarajan: Asymptotic properties of eigenvalues and eigenfunctions of invariant differential operators on symmetric and locally symmetric spaces. 15. G. J. Zuckerman: Quantum physics and semisimple symmetric spaces.

The collection of the research papers published in this and the other two volumes (Lectures Notes in Mathematics, 1024 and 1041) gives an up-to-date account of the areas of the theory of Lie group representations which are of current interest. So this excellent book is highly recommended to everybody who works in this field or on related subjects of mathematics and mathematical physics.

L. Gy. Fehér (Szeged)

Linear and Complex Analysis Problem Book, 199 Research Problems, Edited by V. P. Havin, S. V. Hruščëv and N. K. Nikol'skii (Lecture Notes in Mathematics, 1043), XVIII+719 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984.

The editors write in the Preface: "The most exciting challenge to a mathematician is usually not what he understands, but what still eludes him. This book reports what eluded a rather large group of analysts in 1983 whose interests have a large overlap with those of our Seminar." The 199 problems of the volume derive from more than 200 mathematicians dealing with spectral theory or complex analysis all over the world. They show the main trends of research and the likely directions of further development in these areas. This volume is an extensively expanded version of a former Russian edition which contained 99 problems and was published in 1978. Almost half of these previous problems have been partly or completely solved in the meantime. The solutions are included either as commentaries or in the last "Solutions" chapter. The editors organized the material into the following chapters: 1. Analysis in Function Spaces, 2. Banach Algebras, 3. Probabilistic Problems, 4. Operator Theory, 5. Hankel and Toeplitz Operators, 6. Singular Integrals, BMO, H^p, 7. Spectral Analysis and Synthesis, 8. Approximation and Capacities, 9. Uniqueness, Moments, Normality, 10. Interpolation, Bases, Multipliers, 11. Entire and Subharmonic Functions, 12. C^n, 13. Miscellaneous Problems. Each chapter begins with an introduction, in which, as the editors say, "we try to help the reader to grasp quickly the main point of the chapter, to record additional bibliography, and sometimes also the explain our point of view on the subject or to make historical comments".

This book is a representative collection of problems exciting a wide circle of mathematicians. Every specialist of these fields will certainly read it with great pleasure.

L. Kérchy (Szeged)

D. H. Luecking—L. A. Rubel, Complex Analysis, A Functional Analysis Approach (Universitext), VI+176 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1984.

Let G be an open set on the complex plane and let $H(G)$ denote the set of analytic functions defined in G. $H(G)$ is a topological vector space with the topology of uniform convergence on compact sets. The main point of this book is a representation theorem for the dual space of $H(G)$. The central theorems of complex analysis, like Runge's theorem, Cauchy's integral theorem, Mittag-Leffler's theorem, are derived in an easy way from this result. This "approach via duality is entirely consistent with Cauchy's approach to complex variables, since curvilinear integrals are typical examples of linear functionals". At some places the authors digress even to functions of several variables. At the end of each chapter numerous exercises help the understanding.

This nice book can be a stimulating, delicious reading for every student who is familar with the elements of complex analysis.

L. Kérchy (Szeged)

B. Malgrange, Lectures on the Theory of Functions of Several Complex Variables, 132 pages, Tata Institute Lectures on Mathematics, Bombay, and Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984.

The book is a reissued form of the author's lecture notes published originally in 1958. It provides a very short and well-selected introduction to finite dimensional holomorphic convexity, sheaves, cohomology and Stein manifolds. The treatment is heuristical and it can be considered as almost self-contained: only some quite basic knowledge of holomorphy in one dimension, manifolds with differential forms and locally convex topoligical vector spaces is required for an intelligent reading. The book is divided into three parts: the first focusing on analytic continuation, domains of holomorphy and convexity; the second dealing with d''-cohomology of the cube and the first Cousin problem; the third introducing to the theory of coherent analytic sheaves on Stein manifolds. Each part ends with exercises which give complementary material and with a list of references. There is also a list for supplementary reading but up to date only to 1958.

L. Stachó (Szeged)

J. Marsden—A. Weinstein, Calculus I—II—III (Undergraduate Texts in Mathematics), 1224 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1985.

It is a recurrent care of university lecturers teaching calculus that other sciences (e.g. physics) taught simultaneously to their students require the notions of the derivative and integral very early. In the case of the classical arrangement of topics in a calculus course (real number, limit, continuity, derivation, integration), it is typical that students have to apply notions and methods before they thoroughly get acquainted with them.

Probably it was this challenge that motivated the authors to write this textbook, which is the outgrowth of their experinences while teaching calculus at Berkeley. It is intended for a three-semester sequence with six chapters covered per semesters (four or five student contact hours per week are calculated). At the end of the first semester the students master differentiation and integration and can apply them to solve problems in other sciences. Meanwhile the notion of limit is used intuitively: "If the value of $f(x)$ approximates the number l for x close to x_0, then we say that f approaches the limit l as x approaches x_0." The details about limits involving epsilonics to the different kinds of limits are presented in the second semester, in the context of l'Hospital's rule and infinite series. The differential equations are also found in this semester. The third semester is devoted to functions of vector variables, curves and surfaces, and vector analysis.

The chapters are devided into sections. The first sentence of every section summarizes the main point in italics. Then the notion or method to be introduced is prepared by examples with solutions. The definitions and theorems are placed in emphasizing boxes and are illuminated by further examples and exercises. Every section is concluded by many-many exercises.

The examples and exercises of the book deserve particular attention. Calculus — as the other branches of mathematics — can be picked up only actively, i.e. by solving problems, guessing and proving theorems on one's own way. To do this the students need such well-selected examples and exercises as those in this book. The exercises closing the sections are graded into three consecutive groups: the first exercises are routine, next come those that are still based directly on the examples and the text, the last ones are difficult (these are marked with a star). Answers to odd-numbered exercises are available in the back of the book, and every other odd exercise (i.e. Ex. 1, 5, 9, 13, ...) has a complete solution in the student guide. (The book is supplemented by a Student Guide and an Instructor's Guide.) Anwers to even-numbered exercises are not available to the student. The freshmen who want to know whether or not they posses the prerequisites for reading the book can find several orientation quizzes with answers and a review section to bridge the gap between previous training and the book. The applications show how close calculus is to problems of real world.

As often experienced, it is very difficult to interpret the basic graphs of functions with vector variable and to imagine the complicated surfaces. In this book this is made easier by 1256 figures including plenty of carefully chosen artwork and computer-generated graphics.

This excellent book is highly recommended to every student who wants to learn calculus actively and/or needs it for applications, and to every teacher of calculus who wants to make his classes enjoyable and useful.

L. Hatvani (Szeged)

Tamás Matolcsi, A Concept of Mathematical Physics: Models for Space-Time, 236 pages, Akadémiai Kiadó, Budapest, 1984.

The difficulties physicists encountered in quantum field theory resulted in a crisis of theoretical physics that began in the fifties and lasts up till now. This critical stage of physics bears a strong resemblance to the crisis mathematics faced at the end of the 19th century when a number of paradoxes came to light. The way out from the present troubles of theoretical physics is quite likely to be found in an analogous manner too. What we need is a theoretical physics operating only with notions and only in ways of complete mathematical exactness. As P. A. M. Dirac said: "Any physical or phylosophical ideas that one has must be adjusted to fit the mathematics. Not the way around."

In a critical reaction of mathematical physics, first of all one has to conscientiously analyse even the most common notations. The book under review deals with this task for the space-time notations of physics. It is of primary importance because the entities space-time, matter and field constitute the notational basis present day's physics is built upon.

The book is divided into two parts. Part Ones, Space-Time Models, consists of three chapters. These treat the norrelativistic space-time model, the special relativistic one and the general relativistic space-time models, respectively. Special care is taken for notations like world lines of particles, observers, reference systems and so on. The Galilean, the Lorentz and the Poincaré groups are described in detail. The reader will find useful the exercises given at the end of each chapter of Part One.

The second part of the book is devoted to a concise explanation of the mathematical tools used in the first part. A good account of the material covered here can be given by enumerating the headlines of the chapters in this part: *Tensorial Operations, Pseudo-Euclidean Spaces, Affine Spaces,* Smooth Manifolds, Lie Groups.

Although the exact mathematical formulation of space-time notations of physics constitutes the essence of this work, the physical motivations and interpretations are also sketched by the author. Everyone who is interested in a mathematically clear setting of the basic space-time ideas will find this book useful and enjoyable to read.

L. Gy. Fehér (Szeged)

V. A. Morozov, Methods for Solving Incorrectly Posed Problems, XVIII+257 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1984.

Although the original Russian edition of this book was published ten years ago, it is just as well interesting for mathematicians and engineers today as then.

Several problems of analysis require the solution of the equation (E) $Au=f$, $f\in F$, where $A:D_A\to Q_A$, $D_A\subset U$, $Q_A\subset F$, U and F are metric spaces. The following definition goes back to J. Hadamard: We say that the problem (E) is well-posed if the following three requirements are satisfied: a) Solvability: $Q_A=F$; b) Uniqueness: $Au_1=Au_2$ for any $u_1, u_2\in D_A$ implies $u_1=u_2$; c) Stability: the inverse operator A^{-1} is continuous on F. (This definition was formulated in 1902.) Any mathematical model of a physical problem requires the properties a), b), c). If E does not satisfy all the conditions a), b), c), then this problem is called "ill-posed". Hadamard has constructed an ill-posed problem which became a classical example of the theory. Later many branches of mathematics and natural sciences produced examples involving ill-posed problems, e.g. the continuation for analytic and harmonic functions, automatic control, thermophysics, nuclear physics, the supersonic body problem, biophysical problems etc.

In solving ill-posed problems, the major part of the theory and methods comes from famous Soviet mathematicians. If one has to mention only one name, it is A. N. Tikhonov's. He introduced useful new concepts and discovered several fundamental methods and results. The author of this book also obtained important new results in this field and this fact increases significantly the interest of this book, which in fact discusses mainly the author's investigations. Several theorems can be found in English here for the first time.

L. Pintér (Szeged)

J. D. Murray, Asymptotic Analysis (Applied Mathematical Science, 48), VII+164 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1984.

An earlier version of this book was published by Clarendon Press, Oxford, 1974. Here the very practical Chapter 7 is added on matched asymptotic methods in singular perturbation theory and on multi-scale perturbation methods and suppression of secular terms. The questions discussed in Chapters 1—6 also did not lose their actuality in the past decade. In fact, the domain of possible applications of the methods that are considered in this book becomes wider and wider.

The book is based on lectures given in the mathematics departments at Oxford University and New York University. Chapter 1 contains some necessary definitions, e.g. the definitions of order relations o, O, asymtotic sequences, expansions and series, with many illustrative examples and useful exercises. In the following four chapters the reader can be acquainted with the methods for obtaining analytical approximations, asymptotic expansions to integrals most frequently met in practice, depending on a large or small parameter. In Chapters 6 and 7 the most important asymptotic methods for obtaining asymptotics to solutions of ordinary differential equations with a large or small parameter are given and discussed, for example, the WKB method and the matched expansions. These methods are useful even if the problem posed (strictly speaking) does not have a solution, or if the existence of the solution is not clear.

The table of contents reads as follows. 1. Asymptotic Expansions 2. Laplace's Method for Integrals 3. Method of Steepest Descents 4. Method of Stationary Phase 5. Transform Integrals 6. Differential Equations 7. Singular Perturbation Methods.

The material can easily be covered by undergraduates or graduates with some knowledge of functions of a complex variable and of ordinary differential equations. Many useful examples and exercises are given, there are 25 illustrations and the history of asymptotic methods is also discussed. The book is written in a clear style and without unnecessary details: "Heuristic reasoning, rather than mathematical rigor, is often used to justify a procedure, or some extension of it." It is warmly recommended to everyone interested in differential equations with a large or small parameter and also to other mathematicians or physicists interested in asymptotic expansions, and, finally to every scientists and students interested in mathematical analysis.

J. Hegedűs (Szeged)

Nonlinear Analysis and Optimization, Proceedings of the International Conference held in Bologna, Italy, May 3—7, 1982. Edited by C. Cinti (Lecture Notes in Mathematics, 1107), VI+214 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984.

The conference was organized in honour of Lamberto Cesari, who has recently completed his 70-th year.

In the first lecture D. Graffi talked "On the contributions of Lamberto Cesari to applied mathematics" analysing his works on nonlinear oscillation and nonlinear optics, or more generally, wave propagation in nonlinear media. The second lecture was an interesting essay on the role of applied mathematics and its connection to the physical world, pure mathematics and other sciences. This question has always given rise to much controversy but everybody will probably agree with the following statements of the author, J. Serrin: "... let me add one overarching principle: 'applied' mathematics should be 'good' mathamatics, and whould be marked by the same clarity which all mathematicians necessarily strive for. ... In summary, mathematics can be both necessary and sufficient, bringing order, elegance and beauty to parts of science which otherwise can seem complex, disjoint and confining."

The papers dedicated to L. Cesari suit the subjects of his scientific activity: A. Bensoussan, J. Frehse, Nash point equilibria for variational integrals; H. W. Engl, Behaviour of solutions of nonlinear alternative problems under perturbations of the linear part with rank change; R. P. Gossez, On a property of Orlicz—Sobolev spaces; P. Hess, S. Senn, Another approach to elliptic eigenvalue problems with respect to indefinite weight functions; S. Hildebrandt, Some results on minimal surfaces with free boundaries; R. Kannan, Relaxation methods in nonlinear problems; K. Kirchgässner, Waves in weakly-coupled parabolic systems; J. Mawhin, M. Willem, Variational methods and boundary value problems for vector second order differential equations and applications to the pendulum equation; M. Roseau, Stabilité de régime des machines tournantes et problèmes associés.

In his paper "Nonlinear optimization" L. Cesari gives existence theorems for multidimensional problems of optimal control and for problems of the calculus of variations concerning integrals of an extended Lagrangian on a multidimensional domain.

L. Hatvani (Szeged)

Yasuo Okuyama, Absolute Summability of Fourier Series and Orthogonal Series (Lecture Notes in Mathematics, 1067), VII+18 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984.

With the exception of some classical results such as Bernstein's theorems and some of the authors results from the years 1975—81, the results and problems of the theory of absolute summability of Fourier and orthogonal series considered in this monograph are from the years 1935—1973. Its object is to show that a lot of the classical criteria for absolute convergence can be systematically proved from the point of wiew of best approximation. As absolute summability is a natural extension of absolute convergence, this monograph contains several criteria for the absolute summability of non-absolute convergent Fourier series and orthogonal series. First of all the author collected theorems for absolute Nörlund and Riesz summabilities because absolute Cesaro summability is better known to analysts. This fact is also clear from the contents: Absolute Convergence of Orthogonal Series; Absolute Nörlund Summability Almost Everywhere of Fourier Series; Absolute Nörlund Summability Almost Everywhere of Orthogonal Series; Absolute Riesz Summability Almost Everywhere of Orthogonal Series; Absolute Nörlund Summability Factors of Fourier Series; Absolute Nörlund Summability Factors of Conjugate Series of Fourier Series; Local Property of Absolute Riesz Summability of Fourier Series; Local Property of Absolute Nörlund Summability of Fourier Series.

The reference list contains 96 papers. The book is useful for analysts and graduate students of the field.

I. Szalay (Szeged)

D. P. Parent, Exercises in Number Theory (Problem Books in Mathematics), X+541 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1984.

This book is a translation of the French original published in 1978. D. P. Parent is a pseudonym for 12 French authors, who have collected and chosen these problems in number theory. Many exercises in this book were used for the first time in university examinations. The book is not just a simple exercises-book. There is an introduction to each chapter which contains a summary of the fundamental theorems and notations necessary for the solution of the problems. This theoretical preparation makes the book very useful and gives aid for attempting the solutions. Each chapter is devoted to exactly one area of number theory. Of course it could not have been the aim of the authors that the whole theory be covered in their ten chapters. Solutions are provided for all the problems. The level of these problems, over 150 in number, varies. But these problems have a common feature: every one of them represents an important fact in number theory. This in itself is sufficient to praise the careful work of the authors.

László Megyesi (Szeged)

Probability Theory on Vector Spaces III, Proceedings of a Conference held in Lublin, Poland, August, 24—31, 1983. Edited by D. Szynal and A. Weron (Lecture Notes in Mathematics, 1080), V+373 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984.

Not too long ago, two main characteristics distinguished probability theory as a separate branch of mathematics. One was that this theory used results from almost every other branches of mathematics, the other that this theory rarely contributed to developments in other fields. While the first feature is even more true nowadays, one can assess several streams of a trend opposite to the second.

One particular appearance of this trend is an activity to understand the structure of some abstract vector spaces by means of probabilistic methods. These proceedings fit into this stream, more or less. One of the editors separates three main features of these proceedings. One is the study of stable distributions in abstract settings, represented by four papers. The second one is the investigation of vector-valued processes and Hilbert space methods in stochastic processes, to which topic eight papers are devoted. Weak and strong limit theorems in Hilbert, Orlicz, Banach, or "even" Polish spaces are studied in six papers, three papers deal with ergodic theorems in von Neumann algebras. The remaining four papers are devoted to the study of random functional spaces, elliptically contoured measures, almost sure limits of continuous linear functionals, and to summability questions in Banach lattices.

Those who liked the predecessor Lecture Notes No. 655 and No. 828 will certainly enjoy reading the present continuation.

Sándor Csörgő (Szeged)

R. Remmert, Funktionentheorie I (Grundwissen Mathematik 5), XIII+324 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984.

Who needs yet another book on complex analysis? Nowadays this question seems to be quite reasonable. If the new book has a characteristic feature together with a careful and clear exposition, then we need it. This book possesses these properties. The presentation, though rigorous, is never fussy. In every chapter the reader finds historical remarks. Perhaps this is the most characteristic of this work. For example, reading the interesting history of the Cauchy Integral Theorem one can better understand the difficulties surrounding this result. The Eisenstein's theory of trigonometric functions is interesting for almost everyone. Sometimes we overlook excellent ideas. This book is rich in such ideas.

I found the first volume a very good book and I am looking forward to the second one.

L. Pintér (Szeged)

R. A. Rosenbaum—G. Philip Johnson, Calculus: Basic Concepts and Applications, XVI+422 pages, Cambridge University Press, Cambridge—London—New York—Rochelle—Sidney—Melbourne, 1984.

Various kinds of books exist on calculus. Nevertheless, when we must choose one for our students we can hardly find any suitable.

The cruicial point is that, taking into account the students'age and thorough grounding in mathematics, the spirit of the development must be intuitive with several examples to provide motivation and to clarify concepts only with just a few proofs, yet, the statements must be careful. Such a development may be a good basis of further courses in analysis. In general, in calculus one emphasizes problem solving rather than theory. This is sometimes misleading. If one cannot resist the temptation, the statements will be careless so that the students will have to unlearn many of them later.

This book is a well organized text. The examples, problems are multifarious and instructive. Let us mention one of them proposed at the end of the first chapter (entitled Functional relationships): "The harmonic series is $1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\dots$, a) Calculate $H_n=1+\frac{1}{2}+\dots+\frac{1}{n}$ for $n=1, 2, \dots, 12$. b) Do you think that the H_n will ultimately behave like geometric series with $-1<r<1$, that is, that H_n will get closer and closer to some limiting value as n increases? c) With a programmable calculator or computer, find H_{50}, H_{100} and H_{200}. Also find how large n must be to make $H_n>4$, $H_n>6$ and $H_n>8$. Do these results reinforce or change your guess about the answer to b)? d) You should have found in a) that H_{10} is a little more than 2.9. Note that each of the terms $\frac{1}{11}, \frac{1}{12}, \dots, \frac{1}{99}$

exceeds $\frac{1}{100}$, so $\frac{1}{11}+\frac{1}{12}+\ldots+\frac{1}{100}>\left(\frac{1}{100}\right)90=0.9$. Hence $H_{100}>2.9.+0.9=3.8$. (Actually $H_{100}\approx 5.4$). Similarly the sum of the next 900 terms is greater than 0.9, so $H_{1000}>3.8+0.9=4.7$; and the sum of the next 9000 terms is greater than 0.9, so $H_{1000}>4.7+0.9=5.6$ Use this argument to make a definite statement about H_n as n increases without bound." The authors encourage the students to ask questions constantly: Why is it done this way? Could it not have been accomplished more easily as follows? Is this hypothesis really needed? Does not the following example contradict the statement in the text? How do this problem and its solution compare with other problems I have solved and with situations I know apart from my math course? This as well as the teaching experience striking on every page of the book reminds the reviewer of the outstanding books of Pólya. The authors find the ideal balance between problem solving and theory. The format is pleasing and the printing is excellent. The book is wholeheartedly recommended to students and to teachers who teach calculus.

L. Pintér (Szeged)

Yu. A. Rozanov, Markov Random Fields, IX+201 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1982.

This is the English translation, by Constance M. Elson, of the original Russion edition, Nauka, Moscow, 1980. The present research monograph appears to be the first detailed overview of a relatively new field of research in probability.

When departing from the real line and considering a random field $\{X_t,\ t\in T\subseteqq \mathbf{R}^d\}$, $d\geqq 2$, the notions of "past", "present" and "future", crucial to the Markov property of univariate stochastic processes, loose their unique meaning. There are, therefore, several possible ways to define a "Markov property". The one adopted in this book, so far being the most successful one and dating back to P. Lévy and H. P. McKean, is the following: for each open $S\subseteqq T$ and for each sufficiently small neighbourhood Γ^ε of the boundary Γ of S, the σ-algebras $\sigma\{X_t,\ t\in \bar{S}\}$ and $\sigma\{X_t,\ t\in T\setminus(S\cup\Gamma)\}$ are independent given the σ-algebra $\sigma\{X_t,\ t\in\Gamma^\varepsilon\}$.

Chapter 1 presents necessary technical prerequisites including consistent conditional distributions and Gaussian processes on infinite-dimensional spaces. Chapter 2 starts with the above definition and studies this Markov property in a systematic manner. Chapter 3 is devoted to the study of the Markov property for generalized random functions, i.e., continuous linear mappings of the Schwartz space of infinitely many times differentiable functions on T into the L^2 space on the underlying propbability space, mainly under the assumption of the existence of a dual process. Finally, Chapter 4 gives conditions, in terms of the spectral density, for a vector-valued stationary generalized random field to be Markov.

Sándor Csörgő (Szeged)

N. Z. Shor, Minimization Methods for Non-Differentiable Functions (Springer Series in Computational Mathematics 3) VIII+162 pages, Springer-Verlag Berlin—Heidelberg—New York—Tokyo, 1985.

The main purpose of this book is to give methods for solving nonsmooth optimization problems; it is an English translation ot the original Russian edition. The text starts with an introductory part (Chapter 1) which introduces and investigates special classes of non-differentiable functions, and defines generalized concepts of gradients — namely the concepts of subgradients — either by making use of separability theorems or via the process of taking limits. In Chapter 2 the generalized gradient methods can be found in detailed various versions. Stepsize selection in most of these methods plays a significant role. The relations of these methods to the methods of Fejér-type approximations are shown and the fundamentals of ε-subgradient methods are briefly presented. Chapter 3 is devoted to the description of gradient-type algorithms with space dilation and to the study of the convergence and speed of convergence of these algorithms. Chapter 4 deals with the use of the subgradient methods in iterative algorithms for solving linear and convex programming problems with the aid of some decomposition schemes.

László Gehér (Szeged)

E. Solomon, Games Programming, XI+257 pages, Cambridge University Press, Cambridge—New York—Melbourne, 1984.

"The time is ripe for owners of home computers to raise their sights" — says the author in the Preface.

This is not a book of program listings. The reader is introduced to fundamental concepts of a full utilisation of the machine. References to particular machines are kept to a minimum. The first part of the book deals with those general aspects of computers which are involved in game programming methodology, e.g. program design methods, language processing, program testing, reading play commands, structured data, text handling, random walks, the mathematics of motion, etc. The second part is concerned with simulation games, e.g. war games and management games, where the computer plays the role of the moderator. The third part discusses the implementation of abstract games in which the machine acts as the opponent. Amongs the topics touched are: algorithms and heuristics, game trees, minimax search, the α—β algorithm, self-improving programs.

I. Gyémánt (Szeged)

D. W. Stroock, An Introduction to the Theory of Large Deviations (Universitext), VIII+196 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1984.

These notes are based on a course which the author gave at the University of Colorado. In the Preface he writes: "My intention was to provide both my audience as well as myself with an introduction to the theory of large deviations." He more than achieved this goal and wrote an excellent textbook as well. The first part is devoted to extensions of Cramér's theorem and the second part deals with the theory of large deviations from ergodic phenomena.

Contents: 0. Introduction — 1. Brownian Motion in Small Time, Strassen's Iterated Logarithm — 2. Large Deviations, Some Generalities — 3. Cramér's Theorem — 4. Large Deviation Principle for Diffusion — 5. Introduction to Large Deviations from Ergodic Phenomena — 6. Existence of a Rate Function — 7. Identification of the Rate Function — 8. Some non-Uniform Large Deviation Results — 9. Logarothmic Sobolev Inequalities.

Lajos Horváth (Ottawa and Szeged)

N. M. Swerdlow—O. Neugebauer, Mathematical Astronomy in Copernicus's De Revolutionibus (Studies in the History of Mathematics and Physical Sciences 10), 204 figures, Part 1: XVI+537 pages, Part 2: 538—711 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1984.

Copernicus's "De Revolutionibus Orbium Coelestium" is considered as one of the last, and historically the most important, astronomical work in the Ptolemaic tradition, i.e., using circular orbits to describe the motions of the planets. On the basis of his own and his Greek, Arabic and European predecessors' observations Copernicus's aim was to develop a new heliocentric planetary theory with a moving earth. He attributed three fundamental and a number of secondary motions to the earth.

Except the first eleven chapters of Book I of the "De Revolutionibus" which was devoted to the general description of the universe and of the location and motion of the earth, the whole work consisting of six books is considered in these two volumes. The omission can be explained by the fact that these topics have been treared by several authors.

Part 1 of this book consists of six chapters. The first one is a general introduction containing a detailed biography of Copernicus and an outline of his astronomy. In the remaining five chapters — the headings are: Trigonometry and Spherical Astronomy, The Motion of the Earth, Lunar Theory and Related Subjects, Planetary Theory of Longitude, Planetary Theory of Latitude — we can read a detailed review on Copernicus's book, on his mathematical methods and on his methods of calculation. Part 2 (in a separate volume) contains tables and figures that illustrate the material of Part 1.

This book is unmatched in depth and detail in the literature on Copernicus, and it can be considered as one of the most detailed exposition of a major scientific work ever written.

Piroska Fekete and Lajos Klukovits (Szeged)

A. N. Tikhonov—A. B. Vasil'eva—A. G. Sveshnikov, Differential Equations, VIII+238 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1985.

The theory of differential equations has a special position among the fields of mathematics. It is directly connected to the practice: the models of many evolving, moving systems and processes in the real word are differential equations. So this theory can be considered as a branch of applied mathematics. On the other hand, it has its own life which is independent of the origin of the problems. In this regard it is one of the fields of pure mathematics that raises and solves very deep problems and is in connection with other fields. According to this twofold feature of the theory, the books on differential equations can be either practice-oriented or of a theoretical character. It is not an easy task to write a good book of the first type, especially because it should be readable and understandable with respect to both the contents and the mode of the presentation for users.

Now here is such a book. It is based on a course which has been taught for several years at the Physics Department and the Department of Computational and Cybernetics of Moscow State University. The reader can find interesting physical problems leading to differential equations. Besides the standard topics (existence, uniqueness, dependence of solutions on initial values and parameters, linear equations, boundary value problems, stability theory, first order partial differential equations), two special chapters are included. In one of them we become acquainted with various methods for numerical solution of initial values as well as boundary value problems, and such fundamental notions as the convergence of difference schemes, approximation and stability. The other one gives a brilliant introduction to the asymptotics of solutions of differential equations with respect to a small parameter (in other words, to the theory of regular and singular perturbations).

The style of the book suits its practice-orientad character. If a new idea to be introduced requires complicated techniques, it is presented at first in the possibly simplest case and only after in the general case.

This English translantion can be recommended even to those who have access to the original Russian edition because it includes important improvements.

L. Hatvani (Szeged)

V. S. Varadarajan, Lie Groups, Lie Algebras and their Representations (Graduate Texts in Mathematics, Vol. 102), XIV+430 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1984.

This is a reprint of an introduction to Lie groups, Lie algebras and their representations of wide-ranging popularity which was originally published in the Prentice-Hall Series in Modern Analysis, 1974.

The algebraic as well as the analytic aspects of the theory of Lie groups and their finite-dimensional representations are discussed in detail in the book. The first chapter gives an introductory exposition of the main results of manifold theory that are used throughout the book. In the second chapter, all the basic concepts and results of the general theory of Lie groups and Lie algebras are introduced. To mention some examples: the Lie and the enveloping algebra of a Lie group, the properties of the exponential map, the adjoint representation and the Baker-Campbell-Hausdorff formula are treated here. The third chapter is devoted to the structure theory of Lie algebras. The most important results discussed in this chapter are: the theorems of Lie and Engel on nilpotent and solvable Lie algebras, Cartan's criterion for semisimplicity, Weyl's theorem asserting the semisimplicity of all finite-dimensional representations of a semisimple Lie algebra, and the decomposition theorems of Levi and Mal'cev. The final, most substantial, chapter contains an exchaustive development of the structure and representation theory of semisimple Lie algebras and Lie groups. The classical Lie algebras and the classification of simple Lie algebras over the field of complex numbers are treated here. The representation theory is examined from both the infinitesimal and the global points of view which is an extraordinary merit of this book.

The book under review excels with the clarity of the exposition of its very extensive subject matter which is, otherwise, even widened further by the large number of exercises placed at the end of each chapter. It is a must for everybody interested in the theory of Lie groups, Lie algebras and their representations, from graduate students to specialists and users of the theory.

L. Gy. Fehér (Szeged)

S. Watanabe, Lectures on Stochastic Differential Equations and Malliavin Calculus, 110 pages, Tata Institute of Fundamental Research, Bombay, and Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984.

Consider the partial differential equation (∗) $\partial u/\partial t = Au$, $u(0, x) = f(x)$, where A is a second-order differential operator of the form

$$A = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(x)\frac{\partial}{\partial x_i} + c(x), \quad x = (x_1, \dots, x_d) \in \mathbf{R}^d.$$

In typical cases a stochastic differential equation with respect to a vector-valued Wiener process can be associated with (∗) such that the formal expectation

$$p(t, x, y) = E\Big(\exp\Big\{\int_0^t c(X(s,x))\, ds\Big\}\, \delta_y\big(X(t, x)\big)\Big)$$

where X is the solution of the stochastic equation and δ_y is the Dirac function at $y \in \mathbf{R}^d$, is the fundamental solution of (∗). However, $\delta_y(X(t, x))$ has no meaning as a Wiener functional, a measurable function from a Banach space of continuous vector-valued functions endowed with the supremum norm to $\mathbf{R}^d$. "The purpose of these lectures is to give a correct mathematical meaning to the formal expression $\delta_y(X(t, x))$."

This is a way of presenting Paul Malliavin's infinite-dimensional calculus, introduced in 1976, a stochastic calculus of variations for Wiener functionals. The volume is based on lectures given by the author in 1983 at the Tata Institute, Bangalore, India.

Sándor Csörgő (Szeged)

André Weil, Number Theory: An Approach through History; From Hammurapi to Legendre, XXI+375 pages, Birkhäuser, Boston—Basel—Stuttgart, 1983.

This book is an historical exposition of number theory. The author examines texts that span roughly thirty-six centuries of arithmetical work from an Old Babylonian tablet, detable to the time of Hammurapi to Legendre's Essai sur la Théorie des Nombres (1798). In this very interesting volume A. Weil accompanies the reader into the workshop of four major authors of number theory: Fermat, Euler, Lagrange and Legendre.

Chapter I., the Protohistory, deals with ancient results, e.g., perfect numbers, Pythagorean triangles (partly after the Old Babylonian tablet PLIMPTON 322), indeterminate equations. In Chapter II we can read about Fermat's and his correspondents' works on number theory, e.g., infinite descent, quadratic residues, the prime divisors of sums of two squares. Chapter III deals with Euler's contributions to the subject: large primes, sums of four squares, square roots and continued fractions, Diophantine equations, zeta-function, etc. The last chapter, An Age of Transition: Lagrange and Legendre, deals with indeterminate equations and binary quadratic forms. All the chapters end with appendices in which the reader can find the modern treatments of some problems mentioned in the text, and some detailed original proofs.

While "... enriched by a broad knowledge of intellectual history, Number Theory represents a major contribution to the understanding of our cultural heritage", we recommend this book to the broad mathematical community.

Lajos Klukovits (Szeged)

INDEX

ACTA SCIENTIARUM MATHEMATICARUM

SZEGED (HUNGARIA), ARADI VÉRTANÚK TERE 1

On peut s'abonner à l'entreprise de commerce des livres et journaux „Kultúra" (1061 Budapest, I., Fő utca 32)

ISSN 0234-6523 Acta Univ. Szeged
ISSN 0001-6969 Acta Sci. Math.

INDEX: 26024

86-4610—Szegedi Nyomda—F. v.: Surányi Tibor igazgató

Felelős szerkesztő és kiadó: Leindler László
A kézirat a nyomdába érkezett: 1985. november 21.
Megjelenés: 1986. augusztus

Példányszám: 1 000. Terjedelem: 22 (A/5) ív
Készült monószedéssel, íves magasnyomással,
az MSZ 5601-24 és az MSZ 5602-55 szabvány szerint